

© 2007 by Jordan Eric Vincent. All rights reserved.

QUANTUM MONTE CARLO CALCULATIONS OF THE ELECTRONIC
EXCITATIONS OF GE ATOMS, MOLECULES AND NANOCCLUSERS
USING CORE-POLARIZATION POTENTIALS

BY

JORDAN ERIC VINCENT

B.S., The Ohio State University, 1999

M.S., University of Illinois at Urbana-Champaign, 2003

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2007

Urbana, Illinois

Abstract

In this thesis many-body Quantum Monte Carlo (QMC) calculations are presented for the ground state and excitation energies of Ge atoms, molecules, and clusters as large as $\text{Ge}_{29}\text{H}_{36}$ (≈ 1.2 nm diameter). The hydrogen terminated Ge_nH_m nanoclusters are of particular interest since they have optical properties that are different from small molecules and bulk Ge, and are viewed as an ideal model for quantum confined semiconductor systems. The present QMC results are compared with previous QMC calculations for the corresponding Si molecules and clusters. In addition, the QMC gaps for Ge_nH_m are found to be higher than the gaps reported in recent time-dependent density functional studies, by amounts similar to that previously found for Si systems.

For materials containing Ge it is necessary to deal with the issue of correlation between the core and valence electrons. The accuracy of QMC for heavy atoms such as Ge ($Z=32$) is limited by the fact that the core-valence interactions cannot be treated at the same many-body level as the valence-valence interactions. Typically the core-valence interactions are treated at a single-body level via a pseudopotential, but such methods are unsatisfactory for Ge which has a shallow, easily polarizable $3d$ core. Previous work has proposed using a Hartree-Fock pseudopotential (relativistic) plus a core-polarization potential (CPP) to account for core-valence correlation at a many-body level.

A major part of this work involves quantifying the effect of core-valence correlation via the CPP. The CPP is found to be important for accurate calculations of the total energy, and for excitations of atoms and small molecules. However, there are only small changes in the lowest optical excitations of larger clusters, which can be understood in terms of the nature

of the excitations, mainly the relative Ge s to p atomic orbital character of the excited state.

To my mother and grandmother.

Acknowledgements

The two most important people I want to acknowledge are my mother and grandmother. Without their love and support I would have never been able to complete the work for this thesis.

I want to express my gratitude to all of the great friends I have made in graduate school; all of you helped make life as a physics Ph.D. student in Urbana-Champaign a much more enjoyable experience. I wouldn't have been able to survive here without your help and support.

I especially want to thank my advisor Richard M. Martin for taking me on as his student and helping me complete my thesis. Thank you for giving me so many opportunities and being so helpful and patient throughout my career as a graduate student in your group. You were always very supportive of me and had confidence in my work. I also want to thank Jeongnim Kim for patiently working with me for the past few years and helping me develop and run the codes that I used to do the calculations in this thesis. I have learned a great deal about physics (and C++) in our conversations.

I would also like to thank the following group of people for helping me with my thesis: David Ceperley for advice and help in all aspects of this work, especially regarding quantum Monte Carlo; Eric. L. Shirley for useful discussions pertaining to the CPP formalism and for providing the Ge band structure data; Argyrios Tsolakidis for patiently answering questions regarding the TDLDA work in Ref. [1]; Giulia Galli and Andrew Williamson for useful discussions, especially pertaining to the work on the Si_nH_m nanoclusters in Ref. [2], and for discussions of our QMC results for the Ge_nH_m nanoclusters; Simone Chiesa for help

with basis functions and his role in developing qmcPACK [3]; Nichols Romero and Javier Junquera for help with SIESTA and Dodi Heryadi for help with Gaussian03; Kris Delaney for editing the sections on GW and GW-BSE; Dyutiman Das, Kenneth Esler and Nichols Romero for computer support, many useful conversations and most importantly being great friends.

The work in this thesis was supported by the NSF grant DMR01-04399, the Materials Computation Center grant DMR-0325939 and the DOE Computational Materials Sciences Network through grant DEFG02-91ER45439. Calculations were performed on the NCSA Tungsten, Cobalt and Copper clusters and the CSE Turing cluster.

Table of Contents

List of Tables	xiii
List of Figures	xvi
List of Abbreviations	xix
Chapter 1 Introduction	1
1.1 Thesis Organization	4
1.2 The Born-Oppenheimer Hamiltonian	5
1.3 Exchange and Correlation	6
1.4 A Brief Description of Electronic Excitations	6
1.4.1 Quantum Confinement in Semiconductor Nanoclusters	7
1.4.2 Absorption and Emission in Nanoclusters	7
1.5 Atomic Units	9
1.6 Effective Atomic Units	11
Chapter 2 Theoretical Methods: Single-Body Methods	12
2.1 Introduction	12
2.2 Density Functional Theory	13
2.2.1 Self-Consistency and the Kohn-Sham Equations	15
2.2.2 Approximations to the Exchange-Correlation Functional	15
2.3 The Hartree-Fock Method	16
2.3.1 Introduction and Basic Theory	16
2.3.2 Koopmans' Theorem, Orbital Relaxation and the Delta Self-Consistent Field Method	21
2.4 The Radial Hartree-Fock Equations: Application to Atomic Systems	22
2.4.1 The Spin-Orbitals	22
2.4.2 The Single-Electron Operators	24
2.4.3 Two-Electron Operators: the Hartree Term	25
2.4.4 The Exchange Term	29
2.4.5 Transforming the Non-Local Exchange Operator to a Local Operator	32
2.5 The Scalar-Relativistic Approximation	33
2.5.1 The Dirac Equation	33
2.5.2 Symmetries of the Motion	34
2.5.3 Spin-Orbital Angular Momentum States	35

2.5.4	Eigenvalues of the Spin-Orbital Angular Momentum States	36
2.5.5	Parity and the Operator $\sigma \cdot \hat{\mathbf{r}}$	37
2.5.6	The Operator $\boldsymbol{\sigma} \cdot \mathbf{p}$	38
2.5.7	The Radial Dirac Equation	39
2.5.8	The Koelling-Harmon Approximation	42
2.5.9	Solving the Radial Equations	43
2.6	Results for Germanium	44
2.7	Time-Dependent Density Function Theory	44
2.7.1	The Adiabatic Local Density Approximation	48
2.7.2	Linear Response	48
2.7.3	Explicit Real Time Calculations	49
2.8	Broken Symmetry Solutions	51
Chapter 3	Theoretical Methods: Many-Body Methods	53
3.1	Introduction	53
3.2	Configuration Interaction	53
3.3	The GW Approximation	55
3.4	GW-BSE	57
3.5	Quantum Monte Carlo	58
Chapter 4	Quantum Monte Carlo Methods	59
4.1	Introduction to Monte Carlo	59
4.2	The Central Limit Theorem	61
4.3	Importance Sampling	62
4.4	Markov Chains and the Metropolis Method	63
4.5	Statistics in Monte Carlo	66
4.5.1	Data blocking	67
4.6	Introduction to Quantum Monte Carlo	68
4.7	Variational Monte Carlo	70
4.8	Umbrella or Correlated Sampling	72
4.9	Introduction to Projection Based Methods	74
4.10	Diffusion Monte Carlo	76
4.10.1	The Importance Sampling Transformation	80
4.10.2	Sampling the Drifted Gaussian	82
4.10.3	The Metropolis Accept/Reject Step	83
4.10.4	The Evaluation of Estimators	85
4.11	Diffusion Monte Carlo for Fermion Systems	88
4.11.1	The Fermion Sign Problem	88
4.11.2	The Fixed-Phase and Fixed-Node Approximations	89
4.12	Excited States in QMC	91
4.13	Calculating the Optical Gap in QMC	93

Chapter 5	QMC Algorithms and QMCPack	96
5.1	Introduction	96
5.2	Variational Monte Carlo Algorithm with Importance Sampling for Particle-by-Particle Move	97
5.3	Fixed-Node Diffusion Monte Carlo Algorithm for Particle-by-Particle Move	98
5.4	Parallelization of QMC	102
Chapter 6	The Quantum Monte Carlo Trial Wave Function	104
6.1	Introduction	104
6.2	Evaluating the Slater Determinant	106
6.3	Single-Particle Orbitals	109
6.3.1	Gaussian Type Orbitals and the Gaussian 03 Basis	109
6.3.2	Slater Type Orbitals	111
6.3.3	Numerical Orbitals	112
6.4	The Electron-Electron Jastrow	112
6.4.1	The Derivatives	113
6.4.2	The Ratio	114
6.5	The Electron-Ion Jastrow	115
6.5.1	The Derivatives	116
6.5.2	The Ratio	116
6.6	Examples of Jastrow Functions	117
6.6.1	The Padé Jastrow	117
6.6.2	The Cusp-less Padé-Jastrow	118
6.6.3	The Long-Ranged Jastrow	119
6.7	Evaluating the Local Energy	120
6.8	Optimization of the Trial Wave Function	123
6.8.1	The Cost Function and Correlated Sampling	123
6.8.2	Optimizing the Jastrow	125
6.8.3	Optimization in the Presence of a Non-Local Potential	126
6.8.4	Numerical Instability	127
6.8.5	Outline of the Algorithm	129
Chapter 7	Core-Valence Partitioning	130
7.1	Introduction	130
7.2	Introduction to Pseudopotentials	132
7.2.1	Norm-Conserving Pseudopotentials	133
7.3	Generating Pseudopotentials	135
7.4	The Germanium Pseudopotential	136
7.5	Pseudopotentials in Gaussian	137
7.6	Pseudopotentials in Quantum Monte Carlo	139
7.6.1	Introduction	139
7.6.2	Evaluating Non-Local Pseudopotentials in QMC	139
7.6.3	The Consequences of Using Non-Local Pseudopotentials in Diffusion Monte Carlo	142
7.7	Core Polarization Potentials	144

7.7.1	Core Polarization Potential for Germanium	146
7.7.2	Core Polarization Potentials in QMC: Previous Work	147
Chapter 8 Results for Atomic Systems: Effect of the CPP on Atomic		
	Excitations of Ge	149
8.1	Introduction	149
8.2	Computational Method	150
8.3	Results for Excitations	152
Chapter 9 Results for Molecular Systems		157
9.1	Introduction	157
9.2	Computational Method	157
9.3	Results for Atomization Energies	158
9.4	Treatment of the CPP: Estimating the CPP Correction to Excitations	160
9.5	Results for Excitations	162
9.6	Comparison to Experiment	166
Chapter 10 Results for Nanoclusters		170
10.1	Introduction	170
10.2	The Structure of the Nanoclusters	170
10.3	Previous Experimental and Theoretical Work	173
10.3.1	Review of Experiments	173
10.3.2	Theoretical Overview	174
10.3.3	Focus on Ge Systems	175
10.3.4	Focus on Si Systems	176
10.3.5	The Ge - Si Crossover Question	178
10.4	The Importance of Surface Chemistry	181
10.5	Computational Method	182
10.5.1	The LDA Relaxation of the Structures	183
10.6	Results for the Optical Gaps	186
10.6.1	Comparison of DMC to TDLDA for Ge Clusters	186
10.6.2	Comparison to Previous DMC Results for Si Clusters	189
Chapter 11 Conclusions, General Observations and Future Directions . .		192
Appendix A Appendix		196
A.1	The Spherical Harmonics	196
A.1.1	The Real Spherical Harmonics	198
A.2	Radial Grids	200
A.3	One-Dimensional Cubic Splines	200
A.4	The Runge-Kutta Algorithm	202
A.5	The Numerov Algorithm	204
A.5.1	Application to the Radial Schrodinger Equation and the Need for Log-arithmic Grids	205
A.6	Derivation of the Diffusion Part of the Green's Function	207

A.7	The Cusp Conditions	209
A.7.1	The Nuclear Cusp Condition	209
A.7.2	The Electron-Electron Cusp Condition	210
A.8	Proof of the Relation Between the Norm-Conservation Condition and the Energy Derivative of the Logarithmic Derivative	212
A.9	Input File for Gaussian 03	213
A.10	Data for Germanium	217
A.11	The Local Pseudopotential for Germanium	218
A.11.1	Introduction	218
A.11.2	The Generation of the Pseudopotential	219
A.11.3	Results for Atomic Excitations	221
	References	226
	Author's Biography	239

List of Tables

2.1	Comparison of the results of our radial HF code for scalar-relativistic (SR) and non-relativistic HF (NR) for the ground-state of the Ge atom (3P). We also compare our non-relativistic results to the reference data of Fischer [4]. .	46
4.1	Test results for VMC featuring all-electron move unless otherwise specified. The trial wave function for VMC is identical to the HF wave function and includes no Jastrow correlation factor. All HF wave functions and energies from Ref. [5] unless otherwise specified. All units in a.u. [†] Particle-by-particle move. [‡] HF results from Ref. [6].	72
4.2	The ground-state energy (E^{GS}), excited state energy (E^{EX}) and the energy difference ($\Delta E = E^{GS} - E^{EX} $) for GeH ₄ and Ge ₂ H ₆ from independent VMC simulations compared to the correlated sampling results (VMC-CS). Results with the CPP are also included. All energies in a.u.	75
4.3	The HF, VMC and DMC total energy results for several atoms and molecules compared to experiment. Atomic orbitals from Ref. [5]. All experimental (non-relativistic) results quoted from Ref. [7] and references therein. All energies in a.u. ^a bond length 1.401 Bohr.	87
4.4	Comparison of QMC results from qmcPACK [3] to the published results of Umrigar, Nightingale and Runge (UNR) [8]. For the trial wave function of Ne, see Table V in Ref. [8]; for the trial wave function of Li ₂ , see Table VI in Ref. [8]. The Li ₂ dimer is oriented along the z -axis. The electron-electron Jastrow correlation is the Pade form and is <i>spin-dependent</i>	88
7.1	Comparison of band energy differences for Ge from LDA and quasiparticle results to experiment from Ref [9]. The quasiparticle results feature two different treatments of the core-valence interactions: LDA and a CPP based treatment. The fundamental gaps are in bold. All energies in eV. Experimental results from Ref. [10] unless otherwise noted. [†] From Ref. [9] and references therein.	132
7.2	The spherical grid points and their associated weights: the grid points correspond to the vertices of a regular icosahedron (12 points) on a sphere of radius unity. This is the grid used to evaluate the non-local pseudopotentials in this thesis, including Ge and Si.	142

7.3	Test results for VMC calculations of atoms and molecules featuring pseudopotentials. The trial wave function for VMC is identical to the HF wave function and does not include a Jastrow correlation factor. All HF results are from the Gaussian03 package [11]. The Ge pseudopotential from Ref. [12] is non-relativistic; the atomic basis is the uncontracted set $6s6p1d$. The C pseudopotential and basis are from Refs. [13, 14]. The methane molecule uses the same C pseudopotential and basis as the C atom. The H potential is the bare $-1/r$ and the H basis set $12s1p$ is contracted to $3s1p$. The C-H bond length is 2.06738 Bohr. The second column is the number of core electrons replaced by the pseudopotential (N_{Core}) and the third column is the total number valence electrons ($N_{Valence}$). The reason why the error bar for the Ge calculation is smaller than the calculations involving C is that the Ge pseudopotential has $l_{max} = 2$, while C has $l_{max} = 3$. All energies in a.u.	144
8.1	Comparison of the HF results for the $5s5p1d$ basis to the basis set limit (infinite basis) for the (3P) ground-state of the Ge atom. All energies in a.u. . .	152
8.2	The CPP estimator $\langle \mathcal{H}_{CPP} \rangle$ of various atomic configurations of Ge. All energies in eV.	152
8.3	For the Ge atom results for the $s \rightarrow p$ promotion energy, the s removal energy (s^-), the first four ionization energies and the total energy of the pseudoatom, with and without the CPP. All excitation energies are total energy differences. The CPP correction, ΔCPP , is calculated as the energy difference of the results with and without the CPP. Experimental results from Ref. [15] have been properly spin-orbit corrected [16]: the J weighted averaging has been performed with respect to the $4s^2(^1S)$ configuration, see Appendix A.10. All energies in eV and the QMC error bars are smaller 0.01 eV. GW results from Ref. [17].	155
8.4	Hartree-Fock (HF) and QMC total energies for various atomic configurations of Ge. All energies in eV.	156
9.1	The effect of the CPP (eV) on the ground and excited state energies of Ge_2H_6 calculated by two methods: (i) the differences in the QMC energies with and without the CPP $\langle \Delta E \rangle = \langle \mathcal{H}_{val+CPP} \rangle - \langle \mathcal{H}_{val} \rangle$ and (ii) the CPP estimator $\langle \mathcal{H}_{CPP} \rangle$. The excited state is constructed by promoting a spin-up electron from the HOMO and placing it in the LUMO. Results are also presented for the effect of the CPP on the energy difference between the excited state and ground state, which is the singlet transition.	162
9.2	The QMC results (eV) for the first two singlet transitions of GeH_4 and the first three singlet transitions of Ge_2H_6 , with the corresponding CPP corrections ΔCPP . The transition energies are calculated by Eq. (4.96), where the excited state is created by promoting an electron from the HOMO to the excited state orbital denoted by its irreducible representation. The transition corresponding to the optical gap is in bold. We have used the experimental geometries for GeH_4 and Ge_2H_6	163

9.3	Hartree-Fock (HF) and QMC total energies for the ground-state E^{GS} and excited states E^{EX} of GeH_4 and Ge_2H_6 : the subscript on E^{EX} denotes the unoccupied molecular orbital to which the spin-up electron is promoted. The best available experimental geometries are used. All energies in eV.	166
9.4	The CPP estimator $\langle \mathcal{H}_{CPP} \rangle$ for the ground-state and excited states of GeH_4 and Ge_2H_6 . The best available experimental geometries are used. The excited state is constructed by promoting a spin-up electron from the HOMO and placing it in the MO in parenthesis. All energies in eV.	168
9.5	DMC+CPP results (eV) compared to experiment for the ionization potential (IP) and optical gap (E_{gap}^{opt}) of GeH_4 in its experimental and LDA relaxed geometry. The IP is calculated as the total energy difference of the N and $N - 1$ electron systems: $IP = E[N] - E[N - 1]$. The optical gap is calculated by Eq. 4.96. The experimental data for the IP are from Ref. [18] and the optical gap from [19]. See the text for a discussion of the comparison of our single energy with the two peaks observed in experiment.	169
10.1	QMC optical gaps of the Ge_nH_m , $(n + m) \leq 65$, nanoclusters with the CPP correction ΔCPP . All energies are in eV.	191
A.1	The spherical harmonics $Y_{lm}(\theta, \phi)$ for the s, p, d and f channels.	198
A.2	The real spherical harmonics $S_{lm}(x, y, z)$ in Cartesian coordinates for the s, p, d and f channels. The second column is the index $I_{lm} = l(l + 1) + m$	199
A.3	The energy levels of atomic Ge.	218
A.4	Overall level scheme of energy levels.	219
A.5	Energy levels relative to $4s^2$	219
A.6	Spin-orbit averaged energy levels.	220
A.7	Results from the generation of the local potential: comparison of the eigenvalues from pseudo Hartree-Fock (HF) and all-electron HF.	221
A.8	Comparison of the eigenvalues from the local pseudopotential to the non-local pseudopotential discussed in Section 7.4	221
A.9	The CPP estimator $\langle \mathcal{H}_{CPP} \rangle$ of various atomic configurations of Ge. All energies in eV.	223
A.10	For the Ge atom results for the $s \rightarrow p$ promotion energy, the s removal energy (s^-), the first four ionization energies and the total energy (TE) of the pseudoatom, with and without the CPP. All excitation energies are total energy differences. The CPP correction, ΔCPP , is calculated as the energy difference of the results with and without the CPP. Experimental results from Ref. [15] have been properly spin-orbit corrected [16]: the J weighted averaging has been performed with respect to the $4s^2(^1S)$ configuration [20]. All energies in eV and the QMC error bars are smaller 0.01 eV.	224
A.11	Hartree-Fock (HF) and QMC total energies for various atomic configurations of Ge. For the $(4s)$ configuration, we did not perform QMC since the numerical Hartree-Fock results are exact for a single electron in a central potential. All energies in eV.	225

List of Figures

1.1	Schematic representation of the excitation process. Position (1): the nanocluster is in its electronic ground state and the atomic coordinates of the constituent atoms are in their lowest energy configuration. (1) \rightarrow (2): the nanocluster undergoes a vertical electronic excitation after absorbing a photon. (2) \rightarrow (3): the atomic positions relax to the lowest energy configuration in the excited state. (3) \rightarrow (4): the electron and hole recombine by a vertical transition. Figure originally from Ref. [21].	9
2.1	7.1(a) The radial orbitals for the (3P) ground-state of the Ge atom from all-electron scalar-relativistic Hartree-Fock. 7.1(b) The valence 4s and 4p orbitals.	45
4.1	4.1(a) The results of a VMC simulation for the local energy of an Li_2 dimer. The local energy is computed as an integral over $3N$ -dimensions, where $N = 6$ is the number of electrons. The vertical bar denotes the equilibration length; all averages are computed after this cutoff. 4.1(b) The autocorrelation function of the local energy. The autocorrelation length is approximately equal to 4.0; choosing a block size of this length reduces the correlation between successive values of the local energy to zero.	69
4.2	The results of a Fixed-Node DMC calculation for the Si_5H_{12} cluster. The time-step is 0.025 a.u., the target population is 320 walkers and $n_G = 100$ (see the accompanying text).	95
6.1	Padé-Jastrow $u(r)$ with $a = 1/2$ and several different values of b	118
6.2	Cusp-Less Padé-Jastrow $u(r)$ with $a = 1.0$ and several values of b	119
6.3	The VMC estimator of $\langle T \rangle - \langle \mathbf{F} ^2 \rangle$ for the Li_2 molecule.	122
6.4	Optimization of the ground-state of the Ge atom (3P) using a cost function which is a linear combination of Eqs. (6.71) and (6.72). The optimizable parameters are in the one- and two-electron Jastrow functions; for details regarding the trial wave function see Section 8.2. a) Plots the energy for each step of the optimization and b) plots the variance of the local energy. The optimization algorithm used in this example was conjugate gradient. All units in a.u. . . .	126

6.5	Optimization with constrained weights of the ground-state of the $\text{Ge}_{29}\text{H}_{36}$ nanocluster using the variance as the cost function. The optimizable parameters are in the one- and two-electron Jastrow functions; for details regarding the trial wave function see Section 8.2. a) Plots the energy for each step of the optimization and b) plots the variance of the local energy.	128
7.1	The Ge pseudopotential. The d channel is taken as the local component: the non-local corrections are $(V_s - V_d)$ and $(V_p - V_d)$. All units in a.u.	138
7.2	The V_e term of \mathcal{H}_{CPP} (see Eq. (7.31)) for the Ge atom.	147
8.1	The basis for the ground state (3P) of the Ge atom.	151
9.1	The basis sets for the Ge atoms 9.1(a) and the H atoms 9.1(b). For H the linear combination of 10 Gaussians used to construct the $1s$ and $2s$ orbitals are each fit to a single STO.	159
9.2	The GeH_4 and Ge_2H_6 molecules.	160
9.3	Isosurfaces for several molecular orbitals of GeH_4 , including the HOMO and the first two excited orbitals (red +, blue -). The t_2 MO is triply degenerate. The irreducible representation is also denoted.	164
9.4	Isosurfaces for several molecular orbitals of Ge_2H_6 , including the HOMO and the first three excited orbitals (red +, blue -). The E_u MO is doubly degenerate. The irreducible representation is also denoted.	165
10.1	The Ge_nH_m nanoclusters.	172
10.2	Experimental results for the dependence of photoluminescence spectra with respect to the average diameter of nano-crystalline Ge from Takeoka <i>et al.</i> [22]. The photoluminescence spectrum of the SiO_2 film is also displayed. . .	175
10.3	Difference in the optical gaps for the Si_nH_m nanoclusters predicted by several methods with respect to the DMC results of Williamson <i>et al.</i> [2]: ▲ LDA HOMO-LUMO single particle gaps, Δ LDA quasiparticle gap minus the empirical electron-hole exciton binding energy and corrected to include the electron-hole polarization energy from Ref. [23], □ TDLDA from Ref. [24] (effective gap) ■ TDLDA, ● GW-BSE from Ref. [25].	177
10.4	The lowest-bright (dipole-allowed) singlet excitation energies for several Si_nH_m nanoclusters by Porter <i>et al.</i> [26]. The experimental results for SiH_4 and Si_2H_6 are from Itoh <i>et al.</i> [19] and the results for Si_5H_{12} are from Fehér [27]. The first two peaks in the absorption spectrum are included for SiH_4	179
10.5	Ge_5H_{12} : Isosurfaces for the molecular orbitals used to compute the optical gap (red +, blue -).	183
10.6	$\text{Ge}_{10}\text{H}_{16}$: Isosurfaces for the molecular orbitals used to compute the optical gap.	184
10.7	$\text{Ge}_{14}\text{H}_{20}$: Isosurfaces for the molecular orbitals used to compute the optical gap.	184
10.8	$\text{Ge}_{29}\text{H}_{36}$: Isosurfaces for the molecular orbitals used to compute the optical gap.	185
10.9	Comparison of the present DMC+CPP results for the lowest gaps of the Ge_nH_m , $(n+m) \leq 65$, nanoclusters (●) to the TDLDA results of Nesher <i>et al.</i> [28] (◆). The TDLDA results are “effective gaps” as described in the text. . .	187

10.10	Left: TDLDA results for the absorption spectra of the Ge_nH_m nanoclusters by Nesher <i>et al.</i> [28]. Right: TDLDA results for the absorption spectrum of $\text{Ge}_{29}\text{H}_{36}$ by Tsolakidis and Martin [1], plotted on two energy scales.	188
10.11	Comparison of our DMC results for Ge_nH_m (\bullet) to the DMC results for the corresponding Si nanoclusters by Porter <i>et al.</i> [26] (Δ) and Williamson <i>et al.</i> [2] (∇). We have also included our own DMC results (\circ) for Si_5H_{12} and $\text{Si}_{10}\text{H}_{16}$.	189
A.1	The local Ge pseudopotential. All units in a.u.	222
A.2	Plot of the pseudo and all-electron 4s and 4p Ge atomic orbitals. All units in a.u.	222

List of Abbreviations

au Atomic Units

eV electron Volt

mHartree milli-Hartree

DFT Density Functional Theory

LDA Local Density Approximation

GGA Generalized Gradient Approximation

EXX Exact Exchange

HF Hartree-Fock

DF Dirac-Fock

CI Configuration Interaction

MRCI Multi-Reference Configuration Interaction

CIS Singles-Only Configuration Interaction

CASSCF Complete Active Space Self-Consistent Field Theory

CCSD Coupled-Cluster Singles-Doubles

SCF Self-Consistent Field Theory

Δ SCF Delta Self-Consistent Field Theory

QMC Quantum Monte Carlo

VMC Variational Monte Carlo

DMC Diffusion Monte Carlo

FP Fixed-Phase

FN Fixed-Node
 RMC Reptation Monte Carlo
 PIMC Path Integral Monte Carlo
 TDDFT Time-Dependent Density Functional Theory
 ALDA Adiabatic Local Density Approximation
 TDLDA Time-Dependent Local Density Approximation
 GWA GW Approximation
 BSE Bethe Saltpeter Equation
 EMA Effective Mass Approximation
 TB Tight Binding
 PP Pseudopotential
 NLPP Non-Local Pseudopotential
 NCPP Norm-Conserving Pseudopotential
 AE All-Electron (Full-Core)
 CPP Core Polarization Potential
 H Hamiltonian
 \mathcal{H} Hamiltonian
 \mathcal{H}_{val} Valence-Only Hamiltonian
 \mathcal{H}_{CPP} CPP Hamiltonian
 $\mathcal{H}_{val+CPP}$ Valence-Only Hamiltonian Including CPP
 SM Shirley and Martin
 TM Tsolakidis and Martin
 UNR Umrigar, Nightingale and Runge
 KE Kinetic Energy
 TE Total Energy
 PE Potential Energy
 Σ Self Energy

E_H Hartree Energy
 E_X Exchange Energy
 IP Ionization Potential
 EA Electron Affinity
 E_0 Ground State Energy
 E^{GS} Ground State Energy
 E^{EX} Excited State Energy
 E_T Trial Energy
MO Molecular Orbital
HOMO Highest Occupied Molecular Orbital
LUMO Lowest Unoccupied Molecular Orbital
 Ψ_T Trial Wave Function
 Φ_0 Ground State Wave Function
 Y_{lm} Spherical Harmonic
 S_{lm} Real Spherical Harmonic
 \mathbf{V}_D Drift Velocity
 \mathcal{P} Probability Distribution Function
 $\mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}')$ Transition Probability
 $\mathcal{W}(\mathbf{R} \leftarrow \mathbf{R}')$ Trial Step Probability
 $\mathcal{A}(\mathbf{R} \leftarrow \mathbf{R}')$ Acceptance Probability
 $G(\mathbf{R} \leftarrow \mathbf{R}'; \tau)$ Green's Function
 $G(\mathbf{R}' \leftarrow \mathbf{R}; \tau)$ Backwards Green's Function
 $G_D(\mathbf{R} \leftarrow \mathbf{R}'; \tau)$ Diffusion Green's Function
 $G_B(\mathbf{R} \leftarrow \mathbf{R}'; \tau)$ Branching Green's Function
 $K(\mathbf{R} \leftarrow \mathbf{R}'; \tau)$ Modified Green's Function
 D^\uparrow Spin-Up Slater Determinant
 D^\downarrow Spin-Down Slater Determinant

U_{ee} Electron-Electron Jastrow Correlation

U_{eI} Electron-Ion Jastrow Correlation

\mathbf{R} N-Electron Configuration Vector

E_{gap}^{opt} Optical Gap

PL Photoluminescence

Chapter 1

Introduction

The problem of understanding interacting electrons is one of the most important challenges in science. Electronic correlations control the properties of everyday matter and determine critical properties such as electronic excitations, stability and phases and optical properties [29].

The subject of this thesis research is an accurate treatment of interacting electrons in semiconductor materials such as clusters of silicon and germanium in the nanometer size range using a combination of quantum Monte Carlo (QMC) many-body methods coupled with analytic analysis of the critical issues. Quantum Monte Carlo is ideal because it can be applied to a variety of interacting electron systems and explicitly treats correlation [30, 31, 7, 8]. It is the only known method that can be applied to accurately treat systems of many electrons.

The main focus of the thesis will be on germanium, where the influence of relativity and core-relaxation effects play an integral role in the electronic structure [32, 9]. In addition, Ge is a test case for an important problem, conceptually and practically: the issue of correlation between core and valence electrons. This is not explicitly included in single-body methods and is difficult to include in many-body methods because of the large difference in the energies and spatial extent of core and valence electrons. We have chosen Ge since the effect is largest for shallow, polarizable cores in atoms such as Ge.

As was already alluded to, the proper treatment of the core-electrons for Ge represents a difficulty. Heavy atoms with large cores such as Ge ($Z = 32$) pose a problem for QMC because the computational effort dramatically increases with the atomic number Z [33, 7].

Calculations are greatly simplified by removing the core-electrons and replacing them with an effective-core pseudopotential or pseudo-Hamiltonians [34, 35]; ¹ this thesis deals exclusively with pseudopotentials. Typically pseudopotentials are constructed within single-body methods such as density functional theory (DFT) in the widely used local density approximation (DFT-LDA) or Hartree-Fock (HF), which means that the core-valence interactions of any subsequent QMC calculation on the valence electrons are treated at the same level. While this approach is sufficient for most elements in the periodic table, where core-valence interactions are negligible, it is unsatisfactory for post-transition elements such as Ge which has a shallow and easily polarizable $3d$ core. It is well known that DFT-LDA underestimates the gaps for excitations, whereas HF overestimates the gaps [9, 37, 38, 29]. In fact, Ge is a well known case where the band gap is zero in the DFT-LDA. ²

Core-Polarization potentials [40, 41, 42] (CPP)s represent an improved method to treat the core-valence interactions over mean-field methods such as DFT-LDA and HF. These potentials include dynamical effects of the core dipoles interacting with valence electrons and can be included in all-electron and valence-only calculations. Shirley and Martin (SM) have developed an *ab initio* CPP formalism that corrects the HF treatment of core-valence interactions [43]. Their parameterized CPPs are fit to quasiparticle self-energies, which include many-body effects such as core-relaxation and polarization and core-valence correlation.

QMC calculations for various elements have been performed including CPPs [44, 45, 46, 47, 48, 26], as discussed in Section 7.7. Previous QMC calculations of Ge [49, 50, 51] have not used a CPP and treated the core-valence interactions with local [52] LDA pseudopotentials, which are known to be not as accurate as non-local pseudopotentials, to treat the core-valence interactions. These studies reported results of the total energy the ground-state of

¹An exception are the all-electron QMC calculations for atoms as large as Xe ($Z = 54$) by Towler and co-workers [36]. That work is a tour de force calculation; it is not fully converged and is proposed as an approach that can be feasible since errors in the core may cancel for interesting quantities. This gives support to the idea that it is much better to use pseudopotentials provided it can be shown that they are sufficiently accurate.

²For post-transition elements such as Ge, DFT-LDA overbinds the $4s$ and $4p$ electrons, which is believed to be responsible for underestimating the band gap minimum [9, 39].

the Ge pseudoatom and bands of the solid.

There has been much interest recently in clusters of semiconductors, such as Si and Ge, consisting of tens to hundreds of atoms. Such nanoclusters have properties that can be completely different from small molecules and from bulk materials. In particular, their optical properties have been the focus of intense experimental and theoretical interest due to the observation of visible photoluminescence and their compatibility with existing Si based technology. For Ge nanoclusters, experiments have observed evidence of quantum confinement effects [22, 53, 54, 55] and size dependant photoluminescence in the near-infrared region, which could possibly be due to a radiative recombination of excitons confined in the nanoclusters [22].

The interest in these clusters has led to many theoretical investigations using a variety of methods to predict the effects of quantum confinement on the energy gaps and optical properties. Most of the previous theoretical work has been on Si clusters, including studies using time-dependent density functional theory (TDDFT) [24, 26, 2, 56, 1] and many-body “GW-BSE” [25, 57, 56] and quantum Monte Carlo [26, 2] methods. Theoretical studies for Ge have been much more limited; to our knowledge the only published calculations are ones using tight-binding [58], empirical pseudopotentials [59, 60], density functional theory (DFT) in the local density approximation (LDA) [61, 62], a combination of DFT in the LDA or local spin density approximation (LSDA) with delta self-consistent field theory (“ Δ SCF”) [63, 64, 65] and TDDFT [1, 28]. The qualitative trend of all of these reports is a significant increase in the optical gap as the cluster diameter decreases; however, there are disagreements between the theoretical results.

A major component of this thesis is dedicated to many-body calculations of Ge clusters using the quantum Monte Carlo (QMC) method. To our knowledge there have been no previous QMC calculations for Ge clusters. ³ Results are also reported for atoms and

³During the course of this work it came to our attention that a group at Lawrence Livermore National Laboratory (Andrew Williamson, Aaron Puzder and Giulia Galli) were also doing QMC calculations of the optical gaps of Ge clusters, which closely parallels their work for Si clusters in Ref. [2]. Their results

molecules that test the accuracy of the methods. The effect of core-valence correlation via the CPP is quantified within QMC for total energies and excitation energies. In this thesis only ideal, hydrogen passivated structures are considered. These are the most relevant clusters for the first theoretical study of Ge clusters using QMC, since these are well defined structures and comparisons can be made with previous results for Si clusters having the same structures. However, it should be emphasized that for comparison to experiment, the actual structures are not known, and quantum confinement in ideal structures is not the only mechanism that can affect the optical properties of the nanoclusters. Factors such as the structure and surface chemistry, including surface terminations and reconstructions, have been shown to be important for Si clusters [66, 67, 68, 69, 70]; however, such effects are not considered in the present work.

1.1 Thesis Organization

- This chapter serves as an introduction and includes a brief description of certain fundamental concepts that are used throughout this thesis.
- Chapter 2 discusses various single-body methods such as Hartree-Fock (HF), Density Functional Theory (DFT) and Time-Dependant DFT (TDDFT). Chapter 3 discusses various many-body methods such as Configuration Interaction (CI) and the GW approximation, including GW combined with a solution to the Bethe-Salpeter equation (GW-BSE).
- Chapters 4, 6 and 5 discuss the quantum Monte Carlo (QMC) method and the development of qmcPACK, an new QMC code. Detailed reviews are provided for the Variational Monte Carlo (VMC) and Diffusion Monte Carlo (DMC) methods and the

for Ge clusters are not published and have not been shared with us or the general public at the time of the submission of this thesis. We have discussed our work with them and they have suggested that our independently calculated results are in reasonable agreement with their results.

Fixed-Node (FN) approximation for DMC. Chapter 6 describes the many-body trial wave function used in QMC, and the specific form of the trial wave function that is used for various Ge systems.

- Chapter 7 discusses the concept of core-valence partitioning and the need for pseudopotentials and core-polarization potentials. The specific pseudopotential that is used in place of the Ge atomic core is introduced. In addition, a method of evaluating non-local pseudopotentials (NLPP)s in QMC is reviewed, and discussion is provided for the consequences of using NLPPs within DMC.
- Chapters 8 and 9 quantify the effect of the core-valence correlation on the excitations of atomic Ge (Chapter 8) and the small molecules GeH_4 and Ge_2H_6 (Chapter 9) within QMC. This work is crucial before we move on to our ultimate goal of determining the optical gaps of the hydrogenated Ge nanoclusters Ge_nH_m , for $(n + m) \leq 65$ (Chapter 10).
- In Chapter 10 results for the optical gaps are compared to recent TDLDA results and DMC results for similar Si_nH_m structures are presented. A detailed review of previous experimental and theoretical work on these systems is also provided.

1.2 The Born-Oppenheimer Hamiltonian

The Born-Oppenheimer approximation is based on the assumption that the electronic and ionic degrees of freedom of a system can be decoupled. The N -electron Hamiltonian for the electronic degrees of freedom is given by

$$H = \sum_{i=1}^N \left[-\frac{1}{2} \nabla_i^2 + V_{ext}(\mathbf{r}_i) \right] + \sum_{i < j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (1.1)$$

where \mathbf{r}_i is the coordinate of electron i and V_{ext} is an external potential. The last term is the Coulomb interaction between the electrons. The Born-Oppenheimer Hamiltonian is used throughout this thesis.

1.3 Exchange and Correlation

The wave function of any electronic system must be antisymmetric under the exchange of two electrons due to the exclusion principle for fermions. A consequence of the antisymmetric nature of the wave function is that there is a spatial separation between same spin electrons which reduces the Coulomb energy. This phenomenon is known as the *exchange* interaction; exchange is treated exactly in the Hartree-Fock approximation, see Section 2.3. It is possible to further lower the Coulomb energy by spatially separating opposite spin electrons, but this must be balanced against any increase in the kinetic energy. This phenomenon is known as *correlation*, by definition the *correlation energy* is the difference of the exact energy and the Hartree-Fock energy

$$E_{Corr} = E_{Exact} - E_{HF}. \quad (1.2)$$

The correlation energy is always negative since the Hartree-Fock energy is variational [71].

1.4 A Brief Description of Electronic Excitations

Electronic excitations can be classified into two types: *(i)* single particle excitations in which one electron is removed $N \rightarrow N-1$ or added $N \rightarrow N+1$ to a system, *(ii)* particle conserving or neutral excitations $N \rightarrow N$. The former correspond to electron addition and removals and can be probed experimentally by photoemission or inverse photoemission. Optical properties are determined by the latter class of excitations. Optical excitations are viewed as electron-hole excitations, where an added electron interacts with the hole left by removing electron.

1.4.1 Quantum Confinement in Semiconductor Nanoclusters

For semiconductor nanoclusters quantum confinement, caused by the restricted (nanometric) size of the nanoclusters, can result in superior properties for device applications when compared to those of bulk materials [72]. For Si and Ge nanoclusters, the quantum confinement model predicts a shift in the energies of the valence band and the conduction band edges relative to the bands of bulk; leading to an increased energy gap relative to their bulk counterparts.

One of the most interesting applications of quantum confinement is the ability to manipulate and tune desirable optical properties such as the fundamental absorption and photoluminescence by changing the system size and shape. Size dependant photoluminescence has been observed experimentally in Ge [22] and Si [73, 74] nanoclusters. In addition, a large amount of theoretical work is dedicated to predicting the size dependance of optical properties (gaps) of Si and Ge nanoclusters. Usually these nanoclusters are modeled as bulk-derived structures with the dangling bonds saturated with a passivant such as hydrogen.

Visible photoluminescence has been observed from both Si and Ge nanoclusters produced by a variety of techniques ⁴ [75, 76, 77, 78, 74, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89], a phenomenon not observed at the macroscopic length scale (bulk limit) of either of these materials. (Although there is some controversy over the mechanism responsible for the visible photoluminescence; the role of quantum confinement versus the role of defects). It is not difficult to imagine that visible photoluminescence has a tremendous number of potential applications.

1.4.2 Absorption and Emission in Nanoclusters

Optical *absorption* is the process by which a system is excited from its electronic ground state to an excited state by absorbing light. The absorption of a photon leads to the creation

⁴The most common technique is to grow the nanocrystals in a glass (SiO₂) matrix

of an electron-hole pair (exciton).

When a nanocluster is optically excited to create an electron-hole pair, the final state has approximately the same atomic configuration as the initial state. Hence the excitation is designated as being *vertical*. The pair-excitation energy is the total energy difference between the ground state energy and the energy of the system with the electron-hole pair [21, 65]:

$$E^A = E(N; e + h) - E(N). \quad (1.3)$$

Eq. (1.3) accounts for quasiparticle effects, *i.e.* the reaction of the system to the presence of an extra electron or hole, and the attractive electron-hole Coulomb interaction. The lowest pair-excitation energy describes the onset of optical absorption.

After the optical excitation and prior to *emission* (electron-hole recombination), the coordinates of the constituent atoms relax due to the change of the electronic density, see Figure 1.1. The atoms reach a minimum energy configuration in the presence of the electron-hole pair, which is different than the ground state atomic configuration. The change of the atomic coordinates results in a change in the spectrum; hence the energy levels involved in emission are different than in absorption. This leads to a redshift of the emission lines relative to the absorption lines. The emission energy is defined as the total energy difference

$$E^E = E^*(N; e + h) - E^*(N), \quad (1.4)$$

where (*) designates that the atomic coordinates from the excited system. The Stokes shift is identified as the difference between the onset of optical absorption and the lowest emission energy: $E^A - E^E$. It is assumed that the relaxation of the atoms occurs at a faster rate than the electron-hole recombination [21, 90].

The calculations reported in this thesis are for vertical excitations that correspond to absorption; structural relaxation of the excited state was not performed. Therefore the most appropriate experimental comparison is to photoabsorption measurements. Many experi-

ments measure the emission spectrum which is redshifted with respect to the absorption spectrum.

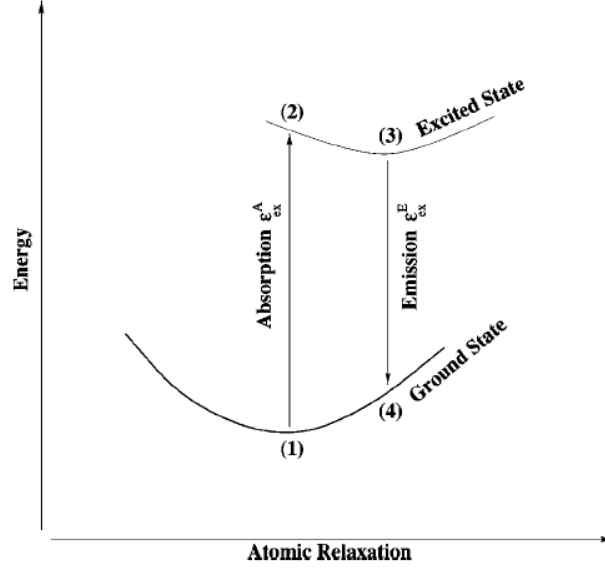


Figure 1.1: Schematic representation of the excitation process. Position (1): the nanocluster is in its electronic ground state and the atomic coordinates of the constituent atoms are in their lowest energy configuration. (1) \rightarrow (2): the nanocluster undergoes a vertical electronic excitation after absorbing a photon. (2) \rightarrow (3): the atomic positions relax to the lowest energy configuration in the excited state. (3) \rightarrow (4): the electron and hole recombine by a vertical transition. Figure originally from Ref. [21].

1.5 Atomic Units

At this point it is necessary to introduce atomic units which are used throughout this thesis. Atomic units are often more convenient to use instead of the International System (SI) of units.

The Schrodinger equation for a single-electron in a Coulomb potential in SI units is

$$\left[-\frac{\hbar^2}{2m_e} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 r} \right] \phi = E\phi. \quad (1.5)$$

To write the Schrodinger equation in dimensionless form make the following change of variables, $\{x, y, z\} \rightarrow \{\lambda x', \lambda y', \lambda z'\}$, where the primed variables are dimensionless. The Schrodinger equation, Eq. (1.5), can be rewritten as

$$\left[-\frac{\hbar^2}{2m_e\lambda^2}\nabla'^2 - \frac{e^2}{4\pi\epsilon_0\lambda r'} \right] \phi' = E\phi'. \quad (1.6)$$

It is possible to solve for λ by the equality

$$\frac{\hbar^2}{m_e\lambda^2} = \frac{e^2}{4\pi\epsilon_0\lambda} = E_a, \quad (1.7)$$

to yield

$$\lambda = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}. \quad (1.8)$$

Divide E by E_a to rewrite the Schrodinger equation in atomic units

$$\left[-\frac{1}{2}\nabla'^2 - \frac{1}{r'} \right] \phi' = E'\phi', \quad (1.9)$$

where $E' = E/E_a$. This is equivalent to setting $\hbar = e^2/(4\pi\epsilon_0) = m_e = 1$. To convert length from atomic units to SI units multiply by the Bohr radius a_0

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.5291772108(18)\text{\AA}, \quad (1.10)$$

to convert energy to SI units multiply by the Hartree E_a :

$$E_a = \frac{m_e e^4}{(4\pi\epsilon_0)^2 \hbar^2} = 27.2113845(23)\text{eV}. \quad (1.11)$$

In atomic units the value of the speed of light in a vacuum is the inverse of the fine

structure constant α :

$$c = \frac{1}{\alpha} \quad \alpha = 7.297352568e - 3 \approx 1/137.04. \quad (1.12)$$

The value of these physical constants are available at Ref. [91].

1.6 Effective Atomic Units

Effective atomic units are mainly useful for quantum dot calculations. The N -electron quantum dot (QD) Hamiltonian can be written as a sum of single-particle terms and an electron-electron Coulomb interaction

$$H = \left[\sum_{i=1}^N -\frac{\hbar^2}{2m^*} \nabla_i^2 + V(\mathbf{r}_i) \right] + \frac{e^2}{4\pi\epsilon^*} \sum_{i < j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (1.13)$$

where m^* is the effective mass and ϵ^* is the dielectric constant for the semiconductor. The effective atomic units are characterized by $\hbar = e^2/(4\pi\epsilon^*) = m^* = 1$. The effective Bohr radius a_o^* and the effective Hartree E_a^* are related to a_0 and E_a :

$$a_o^* = \left(\frac{\epsilon^*}{\epsilon_0} \right) \left(\frac{m_e}{m^*} \right) a_0 \quad E_a^* = \left(\frac{m^*}{m_e} \right) \left(\frac{\epsilon_0}{\epsilon^*} \right) E_a. \quad (1.14)$$

Chapter 2

Theoretical Methods: Single-Body Methods

2.1 Introduction

This chapter is focused on the mean field methods of Hartree-Fock (HF) and Density Functional Theory (DFT) in the Kohn-Sham scheme. These two methods are important because they are the starting point for essentially all correlated-electron calculations. As we will see in Chapter 6, the Quantum Monte Carlo methods uses the results of either HF or DFT to construct a trial wave function.

In HF exchange is treated exactly, the correlation energy is the difference between the HF energy and the exact energy of the many-body system. Hence correlation is neglected in HF. On the other hand, in DFT exchange and correlation are treated approximately by an exchange-correlation functional. See Section 1.3 for a thorough discussion of exchange and correlation.

DFT in the Kohn-Sham scheme is very successful for describing ground-state properties; however, it is far less successful in describing excitations. The Kohn-Sham eigenvalue differences are often used to calculate optical excitations without any formal justifications. In order to properly describe excitations it is necessary to extend DFT to time-dependent systems. The resulting theory is known as time-dependent DFT (TDDFT).

2.2 Density Functional Theory

Density Functional Theory (DFT) [92, 29, 93, 94] is based on the theorems of Hohenberg and Kohn (HK) [95]. The first theorem states that for a system of interacting particles, the ground-state density $n_0(\mathbf{r})$ uniquely determines the external potential V_{ext} up to an additive constant: the inverse relation is also true. A corollary of this theorem is that all the properties of a system are determined by the ground-state density. This is because the density determines the potential, and therefore the Hamiltonian, which determines the many-body wave function for all states (ground and excited) from which physical properties are determined. The second theorem states that the energy of a system is a functional of the density $E[n(\mathbf{r})]$; the functional has its variational minimum at the exact ground-state density. The energy functional is the sum of an interaction term and a term due to the external potential

$$E_{HK}[n] = F_{HK}[n] + \int d\mathbf{r} n(\mathbf{r}) V_{ext}(\mathbf{r}). \quad (2.1)$$

$F_{HK}[n]$ is a *universal functional*, which has no dependence on the external potential $V_{ext}(\mathbf{r})$.

The idea of Kohn and Sham (KS) [96] was to replace the interacting many-body system with an auxiliary independent particle system with the same electron density. The Hohenberg-Kohn energy functional and density can be expressed in terms of single particle orbitals $\{\phi_i\}$. The form of the energy functional in the Kohn-Sham scheme is

$$E[n(\mathbf{r})] = T_0[n(\mathbf{r})] + \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n(\mathbf{r})] + \int d\mathbf{r} n(\mathbf{r}) V_{ext}(\mathbf{r}), \quad (2.2)$$

where $T_0[n]$ is the kinetic energy of the system of density n without electron-electron interactions

$$T_0[n] = \sum_i \int d\mathbf{r} \psi_i^*(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 \right] \psi_i(\mathbf{r}), \quad (2.3)$$

$E_{xc}[n(\mathbf{r})]$ is the exchange-correlation energy and the density is given by

$$n(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2. \quad (2.4)$$

This can be interpreted as a mapping of the interacting many-body system to a system of non-interacting electrons in an effective potential due to the other electrons; the mapping is exact if $E_{xc}[n]$ is known exactly. The functional in Eq. (2.2) is minimized, with the orthonormality constraint $\int d\mathbf{r} \psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) = \delta_{ij}$, to produce the set of Kohn-Sham equations

$$\left[-\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \quad (2.5)$$

where $V_H(\mathbf{r})$ is the Hartree potential

$$V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (2.6)$$

and $V_{xc}(\mathbf{r})$ is the exchange-correlation potential. More formally, the exchange-correlation potential is the variational derivative of the exchange-correlation energy

$$V_{xc}(\mathbf{r}) = \left. \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \right|_{n(\mathbf{r})=n_0(\mathbf{r})}. \quad (2.7)$$

The many-body terms are incorporated into the exchange-correlation functional of the density.

It is important to note that the single particle energies $\{\epsilon_i\}$ obtained from the Kohn-Sham equations do not have any physical meaning; there is also no counterpart in DFT to Koopmans' theorem, see Section 2.3.2.

2.2.1 Self-Consistency and the Kohn-Sham Equations

The Kohn-Sham equations, Eq. 2.5, are solved iteratively. As a first step an initial guess is made for the density n , this is the zeroth iteration. The density is used to construct the corresponding effective potential: $V_{eff}(\mathbf{r}) = V_{ext}(\mathbf{r}) + V_H[n] + V_{xc}[n]$. The Kohn-Sham equations are solved for the effective potential to obtain a new density, which leads to a new effective potential. This process is repeated until *self-consistently*. The effective potential determined by this method is known as the self-consistent or mean field.

2.2.2 Approximations to the Exchange-Correlation Functional

The exchange-correlation energy $E_{xc}[n(\mathbf{r})]$ contains Coulomb correlation effects beyond the Hartree approximation (the V_H term) and part of the kinetic energy for the interacting system ($T - T_0$), where T is the interacting kinetic energy. For real systems $E_{xc}[n(\mathbf{r})]$ is not known: the most widely used approximation for $E_{xc}[n(\mathbf{r})]$ is the local density approximation (LDA). The LDA uses exact expressions for the exchange energy and various interpolation schemes to numerically exact quantum Monte Carlo (QMC) correlation energies for the homogeneous electron gas [97, 98]

$$E_{xc}^{LDA}[n] = \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}^{hom}(n(\mathbf{r})), \quad (2.8)$$

where $\epsilon_{xc}^{hom}(n)$ is the exchange-correlation energy per electron of a uniform electron gas of density n . [96] The generalization to spin-polarized systems, where the total energy is a functional of the spin-up n^\uparrow and spin-down densities n^\downarrow , is the local spin density approximation (LSDA) [93]. The LDA can be applied to inhomogeneous systems, where it has been shown to give a qualitatively good description of the ground-state properties. However, it is well known that the accuracy of DFT-LDA is limited and almost always overestimates the binding energy and underestimates the bond lengths of molecules and solids [93, 29]. In addition, one of the faults of LDA and LSDA is that the self-interaction terms are only approximately

canceled, in contrast the self-interaction is exactly canceled in the HF approximation.

The generalized gradient approximations (GGA) [92, 99, 99, 100, 101], which include gradient corrections to the local density, are an improvement to LDA. Exact exchange (EXX) [102] is a way to include the correct exchange within the Kohn-Sham DFT formalism.

2.3 The Hartree-Fock Method

2.3.1 Introduction and Basic Theory

The Hartree-Fock method [103, 104, 16, 105, 4, 106, 107, 108, 29] is an independent-particle approximation to solve the N -electron system with the non-relativistic Hamiltonian

$$H = \left(\sum_{i=1}^N -\frac{1}{2} \nabla_i^2 + V_{ext}(\mathbf{r}_i) \right) + \frac{1}{2} \sum_{i \neq j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.9)$$

where V_{ext} is the external potential. For atoms and molecules $V_{ext} = -Z/|\mathbf{r} - \mathbf{R}_\alpha|$ unless the core-electrons are replaced by a pseudopotential. This Hamiltonian can be written in the form

$$H = H_1 + H_2, \quad (2.10)$$

where H_1 is the sum of N single-electron terms

$$H_1 = \sum_i h(i), \quad h(i) = -\frac{1}{2} \nabla_i^2 + V_{ext}(\mathbf{r}_i) \quad (2.11)$$

which consists of the kinetic energy of a single-electron and the interaction of a single-electron with the external potential, and H_2 is the sum of $N(N-1)/2$ two-electron terms

$$H_2 = \sum_{i < j} g(i, j), \quad g(i, j) = \frac{1}{r_{ij}} \quad (2.12)$$

which consists of the electron-electron interactions (with $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$).

In Hartree-Fock the electrons are treated as non-interacting with the exception that they must obey the fermion exclusion principle; correlations are treated at a Hartree interaction level and not in a true many-body sense. The method uses as a trial wave function a single Slater determinant of spin-orbitals

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix}, \quad (2.13)$$

where $\mathbf{x}_i = \{\mathbf{r}_i, \sigma_i\}$ represents the space and spin coordinates of the i th electron. This trial wave function automatically obeys fermion particle exchange anti-symmetry

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N) = -\Psi(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N), \quad (2.14)$$

interchanging any row or column of the Slater determinant results in a sign change. The spin-orbitals are constrained to be orthonormal

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} \quad (2.15)$$

and can be represented as a product of a spatial orbital $\psi_i(\mathbf{r})$ and a spin function (α for spin-up or β for spin-down). Since the spin-up spin-orbitals are automatically orthogonal to the spin-down spin-orbitals, the condition in Eq. 2.15 reduces to the condition that the spatial part of the spin-orbitals with the same spin are orthonormal. From Eq. (2.15), it follows that Ψ is normalized, $\langle \Psi | \Psi \rangle$, since the spin-orbitals that form the matrix elements of the Slater determinant are orthonormal.

The expectation value of the energy is given by the sum of the expectation values of the

single-electron operators and the two-electron operators:

$$\langle H \rangle = \langle H_1 \rangle + \langle H_2 \rangle, \quad (2.16)$$

where $\langle \rangle$ denotes the average with respect to Ψ . The expression for the single-electron operators is given by

$$\langle \Psi | H_1 | \Psi \rangle = \sum_i I_i, \quad (2.17)$$

where

$$\begin{aligned} I_i &= \langle \psi_i(\mathbf{x}_1) | h_1 | \psi_i(\mathbf{x}_1) \rangle \\ &= \int d\mathbf{r}_1 \psi_i^*(\mathbf{r}_1) h_1 \psi_i(\mathbf{r}_1). \end{aligned} \quad (2.18)$$

The expression for the two-electron operators consists of the sum of the *direct* and *exchange* terms

$$\langle \Psi | H_2 | \Psi \rangle = \frac{1}{2} \sum_i \sum_j [J_{ij} - K_{ij}]. \quad (2.19)$$

The direct term

$$\begin{aligned} J_{ij} &= \langle \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) | \frac{1}{r_{12}} | \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) \rangle \\ &= \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_i^*(\mathbf{r}_1) \psi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \psi_i(\mathbf{r}_1) \psi_j(\mathbf{r}_2), \end{aligned} \quad (2.20)$$

is the average value of the interaction $1/r_{12}$ when electron 1 is in the spin-orbital ψ_i and electron 2 is in the spin-orbital ψ_j . The exchange term

$$\begin{aligned} K_{ij} &= \langle \psi_i(\mathbf{x}_1) \psi_j(\mathbf{x}_2) | \frac{1}{r_{12}} | \psi_j(\mathbf{x}_1) \psi_i(\mathbf{x}_2) \rangle \\ &= \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_i^*(\mathbf{r}_1) \psi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \psi_j(\mathbf{r}_1) \psi_i(\mathbf{r}_2), \end{aligned} \quad (2.21)$$

is the matrix element of the interaction $1/r_{12}$ with respect to $\psi_i(\mathbf{r}_1) \psi_j(\mathbf{r}_2)$ and $\psi_j(\mathbf{r}_1) \psi_i(\mathbf{r}_2)$

- electron 1 and electron 2 are interchanged.

The Hartree energy is commonly written as

$$E_H = \frac{1}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (2.22)$$

which represents the self-interaction of the density $n(\mathbf{r})$: this is the reason why the $i = j$ term is included in the double summation. The self-interaction in the Hartree term is canceled by a corresponding term in the exchange. It is important to note that the exchange term only contributes for same spin electrons; this is a direct result of orthogonality of the spin-orbitals.

Using Eqs. (2.17, 2.18, 2.19, 2.20, 2.21), the expectation value of the Hamiltonian, Eq. (2.9), with respect to the Slater determinant trial wave function can be written as:

$$\begin{aligned} E(\{\Psi\}) = \langle \Psi | \hat{H} | \Psi \rangle &= \int_0^\infty d\mathbf{r} \psi_i^*(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + V_{ext}(r) \right] \psi_i(\mathbf{r}) \\ &+ \frac{1}{2} \sum_{i,j} \int d\mathbf{r}' d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}') \psi_i(\mathbf{r}) \\ &- \frac{1}{2} \sum_{i,j} \int d\mathbf{r}' d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}) \psi_i(\mathbf{r}'). \end{aligned} \quad (2.23)$$

The sum over the spin-orbitals runs over all the occupied spin-orbitals, we have not explicitly included any occupation coefficients.

The energy $E[\Psi]$ is a functional of the Slater determinant, or the constituent spin-orbitals, and is variational: $E_0 \leq E[\Psi]$, where E_0 is the exact ground-state energy of the system given by the Hamiltonian in Eq. (2.9). The goal is to find the minimum of this functional with respect to the spin-orbitals: the spin-orbitals corresponding to this minimum are occupied to construct the ground-state of the system. The minimization must be performed with the orthonormality constraint of the spin-orbitals in Eq. (2.15); the result being the set of

equations

$$\begin{aligned} \left(-\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r})\right) \psi_i(\mathbf{r}) &+ \sum_j \int d\mathbf{r}' \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}) \\ &- \sum_j \delta_{\sigma_i, \sigma_j} \int d\mathbf{r}' \frac{\psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \end{aligned} \quad (2.24)$$

where the sum j is over all the occupied spin-orbitals and ϵ_i is a Lagrange multiplier. A more compact form of Eq. (2.24) is given by

$$\left[-\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + \hat{V}_H + \hat{V}_X\right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (2.25)$$

where we have introduced the direct operator

$$\hat{V}_H \psi_i(\mathbf{r}) = \int d\mathbf{r}' \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}) \quad (2.26)$$

and the non-local exchange operator

$$\hat{V}_X \psi_i(\mathbf{r}) = - \sum_j \delta_{\sigma_i, \sigma_j} \int d\mathbf{r}' \frac{\psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}). \quad (2.27)$$

The Hartree-Fock equations can now be written in a form analogous to the Schrodinger equation with an effective potential

$$\left[-\frac{1}{2}\nabla^2 + \hat{V}_{eff}^i\right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (2.28)$$

where

$$\hat{V}_{eff}^i = V_{ext}(\mathbf{r}) + \hat{V}_H^i + \hat{V}_X^i. \quad (2.29)$$

It is important to note that \hat{V}_{eff}^i explicitly depends on the spin-orbital ψ_i and is a non-local operator, hence the labeling.

The HF equations, Eq. (2.24), are solved self-consistently in a process that is very similar to the method to solve the Kohn-Sham equations. The HF equations are typically either solved in a *spin-restricted* or *spin-unrestricted* fashion. For spin-restricted systems, each of the spin-orbitals is doubly occupied by one spin-up electron and one spin-down electron. The ground-state is constructed by filling the first $N/2$ spin-orbitals. For spin-unrestricted HF, the spin-orbitals are allowed to have an explicit spin dependence; each spin-orbital is only singly occupied. Since the spin-unrestricted version of HF has more variational freedom than the spin-restricted version, the total energy from an unrestricted HF calculation is lower than the total energy of a restricted HF calculation.

Taking the scalar product of Eq. (2.24) with $\psi_i(\mathbf{r})$ and using Eqs. (2.18, 2.20, 2.21), it can be shown that the single particle energy is given by the expression

$$\epsilon_i = I_i + \sum_j [J_{ij} - K_{ij}]. \quad (2.30)$$

This relation implies that the total Hartree-Fock energy can also be given by

$$E[\Psi] = \sum_i \epsilon_i - \frac{1}{2} \sum_i \sum_j [J_{ij} - K_{ij}]. \quad (2.31)$$

An important observation is that the total energy is *not* the sum of the individual eigenenergies - the electron-electron interaction energies are double counted.

2.3.2 Koopmans' Theorem, Orbital Relaxation and the Delta Self-Consistent Field Method

Koopmans' theorem states that the eigenvalues from the HF equations can be identified as total energy differences. For example, the energy to remove an electron occupying the orbital ψ_i corresponds to the eigenvalue ϵ_i . Koopmans' theorem is based on the assumption that the single-particle orbitals are frozen upon adding or removing an electron. In reality the

orbitals change when an electron or hole, or an electron-hole pair, is added to a system. This phenomenon is known as *relaxation*. It is possible to include the effects of orbital relaxation by the delta self-consistent field (Δ SCF) method. In the Δ SCF approximation, addition and removal energies are calculated as total energy differences, *i.e.* the energy difference of two SCF calculations. Weissker and co-workers [63, 64, 65] have used the Δ SCF method to calculate the optical properties of the Si_nH_m and Ge_nH_m nanoclusters.

2.4 The Radial Hartree-Fock Equations: Application to Atomic Systems

From Section 2.3.1 the Hartree-Fock equations, Eq. (2.24), can be written in terms of the spatial part of the spin-orbitals. For atomic systems, or any system with a spherically symmetric central potential, it is possible to solve the set of Hartree-Fock equations numerically on a radial grid by using the Numerov algorithm, see Section A.5. The integrals for the matrix elements of the Hamiltonian are solved by a method discussed in Appendix A.4. In general it is necessary to use a basis, such as Gaussian or Slater type orbitals, to solve the Hartree-Fock equations.

The derivation of the radial HF equations closely follows that of Refs. [16, 105]. As part of this thesis we have written a numerical HF code [109] for atoms and quantum dot systems with a harmonic central potential. The results from this code represent the HF basis set limit: the result of a HF calculation with an infinite basis.

2.4.1 The Spin-Orbitals

For an atomic system the spin-orbital ψ_i index i represents the set of four quantum numbers (n_i, l_i, m_i, s_i) . For spherically symmetric external potentials the spatial part of a spin-orbital

can be written as a product of a radial function and a spherical harmonic

$$\psi(\mathbf{r}) = R_{nl}(r)Y_{lm}(\theta, \phi) = \frac{u_{nl}(r)}{r}Y_{lm}(\theta, \phi). \quad (2.32)$$

The spherical harmonic is characterized by the quantum numbers (l, m) (for a detailed discussion of the spherical harmonics see Appendix A.1). In general the radial component $R_{nl}(r)$ can be characterized by the quantum numbers (n, l, m, σ) - a more detailed discussion of this is given later in this section.

The radial function obeys the normalization condition

$$\int_0^\infty dr r^2 |R_{nl}(r)|^2 = \int_0^\infty dr |u_{nl}(r)|^2 = 1. \quad (2.33)$$

Applying the new representation of the spin-orbital to the radial Schrodinger equation for a potential $V(r)$

$$\left[-\frac{1}{2}\nabla^2 + V(r) \right] \psi_{nlm}(\mathbf{r}) = \epsilon_{nlm} \psi_{nlm}(\mathbf{r}), \quad (2.34)$$

where the Laplacian in spherical coordinates is given by

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}, \quad (2.35)$$

leads to ($R = u/r, R' = u'/r - u/r^2, R'' = u''/r - 2u'/r + 2u/r^3$)

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V(r) \right] u_{nl}(r) = \epsilon_{nl} u_{nl}(r). \quad (2.36)$$

The analytic solution for the eigenvalue for the one-electron atom with the central potential $V(r) = -Z/r$ is given by the expression

$$\epsilon_{nl} = -\frac{Z^2}{2n^2}, \quad (2.37)$$

where n is the principal quantum number. The restriction on the angular momentum quantum number is $0 \leq l < n$. The solution to the radial functions $R_{nl}(r)$ are related to the Laguerre polynomials and can be found in Refs [110]. The nodes of the radial functions are determined by $n - l - 1$, with the $s(l = 0)$ states having the most nodes. The nodes are intimately related to the orthogonality constraint.

For *closed-shell* systems it can be shown that the charge density $n(\mathbf{r})$ has spherical symmetry

$$n(\mathbf{r}) = \sum_i^{\text{occupied}} |\psi_i(\mathbf{r})|^2 = \sum_{nlm\sigma}^{\text{occupied}} \frac{|u_{nl}(r)Y_{lm}(\theta, \phi)|^2}{r^2} = \sum_{n,l}^{\text{occupied}} \frac{2(2l+1)}{4\pi} \frac{|u_{nl}(r)|^2}{r^2}, \quad (2.38)$$

by use the relation in Eq. (A.10). Closed-shell means that each spin-up electron is paired with a spin-down electron. From this observation it is not difficult to show that the effective potential also is spherically symmetric. For a closed-shell system the radial orbitals $u(r)$ only depend on the quantum numbers (n, l) . For an open-shell system, such as the ground-state of the Ge atom, the radial function can depend on all of the quantum numbers (n, l, m, σ) . It is possible to perform a *restricted* HF calculation where orbitals with the same (n, l, m) or (n, l) quantum numbers are forced to be identical.

2.4.2 The Single-Electron Operators

The single-electron operators include the Laplacian, Eq. (2.36), and the external potential. Evaluating the matrix elements of the single-electrons operators poses a challenge due to the second derivative introduced by the Laplacian. The evaluation of the second derivatives can be avoided by noting

$$\int_0^\infty d\mathbf{r} \psi_i^*(\mathbf{r}) \left[-\frac{1}{2}\nabla^2 + V_{ext}(r) \right] \psi_i(\mathbf{r}) = \int_0^\infty d\mathbf{r} \psi_i^*(\mathbf{r}) [\epsilon_i - (V_{eff}^i - V_{ext}(r))] \psi_i(\mathbf{r}), \quad (2.39)$$

where ϵ_i is the eigenvalue. This allows the total energy to be written as

$$E = \sum_i \epsilon_i - E_H - E_X. \quad (2.40)$$

2.4.3 Two-Electron Operators: the Hartree Term

The Potential

The Hartree operator acting on the i th orbital

$$\hat{V}_H \psi_i(\mathbf{r}) = \left[\sum_j \int d\mathbf{r}' \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \right] \psi_i(\mathbf{r}), \quad (2.41)$$

can be written in terms of the radial functions $u_i(r) = rR_i(r)$, Eq. (2.32), with help of the following two identities:

1. The multi-pole expansion

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{r_{<}^k}{r_{>}^{k+1}} \frac{4\pi}{2k+1} Y_{kq}(\theta, \phi) [Y_{kq}(\theta', \phi')]^*, \quad (2.42)$$

where $r_{<}$ is the smaller and $r_{>}$ is the larger of r_1 and r_2 .¹

2. The integral relation among the spherical harmonics

$$\int d\Omega [Y_{l_3, m_3}(\theta, \phi)]^* Y_{l_1, m_1}(\theta, \phi) Y_{l_2, m_2}(\theta, \phi) = \sqrt{\frac{(2l_1+1)(2l_1+1)}{4\pi(2l_3+1)}} \langle l_1 m_1 l_2 m_2 | l_3 m_3 \rangle \langle l_1 0 l_2 0 | l_3 0 \rangle. \quad (2.43)$$

¹See Eq. (3.70) on pg. 102 of [111]

which obeys the following conditions:

$$\begin{aligned}
m_1 + m_2 &= m_3 \\
l_1 + l_2 + l_3 &= \text{even, else } 0 \\
l_1 + l_2 &\geq l_3 \geq |l_1 - l_2|.
\end{aligned} \tag{2.44}$$

Using identities 1. and 2., the Hartree term takes the form

$$\begin{aligned}
\hat{V}_H \psi_i(\mathbf{r}) &= \sum_j \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{4\pi}{2k+1} Y_{kq}(\theta, \phi) \overbrace{\int_0^{\infty} dr' u_{n_j l_j}^*(r') u_{n_j l_j}(r') \frac{r_{<}^k}{r_{>}^{k+1}}} \times \\
&\left[\int d\Omega' Y_{l_j, m_j}(\theta', \phi') [Y_{l_j, m_j}(\theta', \phi')]^* [Y_{kq}(\theta', \phi')]^* \right] \frac{u_{n_i l_i}(r)}{r} Y_{l_i, m_i}(\theta, \phi). \tag{2.45}
\end{aligned}$$

The most generic form of the radial integral over the radial functions from the previous equation

$$\int_0^{\infty} dr' u_{n_i l_i}^*(r') u_{n_j l_j}(r') \frac{r_{<}^k}{r_{>}^{k+1}} \tag{2.46}$$

can be written in terms of the function introduced by Slater ²

$$\frac{\mathcal{Y}_k(n_i l_i; n_j l_j / r)}{r} = \frac{1}{r^{k+1}} \int_0^r dr' r'^k u_{n_i l_i}^*(r') u_{n_j l_j}(r') + r^k \int_r^{\infty} dr' r'^{-k-1} u_{n_i l_i}^*(r') u_{n_j l_j}(r') \tag{2.47}$$

Now it is possible to write the Hartree term as

$$\begin{aligned}
\hat{V}_H \psi_i(\mathbf{r}) &= \sum_j \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{4\pi}{2k+1} Y_{kq}(\theta, \phi) Y_{l_i, m_i}(\theta, \phi) \frac{\mathcal{Y}_k(n_j l_j, n_j l_j / r)}{r} \frac{u_{n_i l_i}(r)}{r} \times \\
&\sqrt{\frac{(2k+1)(2l_j+1)}{4\pi(2l_j+1)}} \langle l_j m_j k q | l_j m_j \rangle \langle l_j 0 k 0 | l_j 0 \rangle
\end{aligned}$$

²See Eq. (17-27) on pg. 17 of [105]

with the restrictions

$$\begin{aligned}
m_j + q &= m_j \implies q = 0 \\
l_j + k + l_j &= \text{even, else } 0 \\
l_j + k &\geq l_j \geq |l_j - k| \implies 0 \leq k \leq 2l_j.
\end{aligned} \tag{2.48}$$

To eliminate any dependence on θ or ϕ perform the integration

$$\int d\Omega [Y_{l_i, m_i}(\theta, \phi)]^* \left[\sum_j \int d\mathbf{r}' \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}) \right]. \tag{2.49}$$

which leads to

$$\begin{aligned}
&= \sum_j \sum_{k=0}^{\infty} \sum_{q=-k}^k \sqrt{\frac{4\pi}{2k+1}} \frac{\mathcal{Y}_k(n_j l_j; n_j l_j / r)}{r} \frac{u_{n_i l_i}(r)}{r} \times \\
&\quad \langle l_j m_j k q | l_j m_j \rangle \langle l_j 0 k 0 | l_j 0 \rangle \int d\Omega [Y_{l_i, m_i}(\theta, \phi)]^* Y_{kq}(\theta, \phi) Y_{l_i, m_i}(\theta, \phi) \\
&= \sum_j \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{4\pi}{2k+1} \frac{\mathcal{Y}_k(n_j l_j; n_j l_j / r)}{r^2} u_{n_i l_i}(r) \times \\
&\quad \langle l_j m_j k q | l_j m_j \rangle \langle l_j 0 k 0 | l_j 0 \rangle \langle l_i m_i k q | l_i m_i \rangle \langle l_i 0 k 0 | l_i 0 \rangle
\end{aligned}$$

with the additional restrictions

$$\begin{aligned}
m_i + q &= m_i \implies q = 0 \\
l_i + k + l_i &= \text{even, else } 0 \\
l_i + k &\geq l_i \geq |l_i - k| \implies 0 \leq k \leq 2l_i.
\end{aligned} \tag{2.50}$$

Utilizing the symmetry properties of the Clebsch-Gordan coefficients [112]

$$\langle l_1 m_1 l_2 m_2 | l_3 m_3 \rangle = \langle l_2 - m_2 l_1 - m_1 | l_3 - m_3 \rangle \quad (2.51)$$

$$= (-1)^{l_1+l_2+l_3} \langle l_2 m_2 l_1 m_1 | l_3 m_3 \rangle \quad (2.52)$$

$$= (-1)^{l_1+l_2+l_3} \langle l_1 - m_1 l_2 - m_2 | l_3 - m_3 \rangle \quad (2.53)$$

$$= (-1)^{l_1-m_1} \sqrt{\frac{(2l_3+1)}{2l_2+1}} \langle l_1 m_1 l_3 - m_3 | l_2 - m_2 \rangle \quad (2.54)$$

$$= (-1)^{l_2-m_2} \sqrt{\frac{(2l_3+1)}{2l_1+1}} \langle l_3 - m_3 l_2 m_2 | l_1 - m_1 \rangle \quad (2.55)$$

and properly taking into account the restrictions, it is now possible to write the final solution

$$\begin{aligned} \hat{V}_H \frac{u_{n_i l_i}(r)}{r} &= \sum_j \sum_{k=0}^{\min(2l_i, 2l_j)} (-1)^{-m_i-m_j} \frac{(2l_j+1)(2l_i+1)}{(2k+1)^2} \langle l_j m_j l_j (-m_j) | k0 \rangle \langle l_j 0 l_j 0 | k0 \rangle \times \\ &\quad \langle l_i m_i l_i (-m_i) | k0 \rangle \langle l_i 0 l_i 0 | k0 \rangle \frac{\mathcal{Y}_k(n_j l_j; n_j l_j/r)}{r} \frac{u_{n_i l_i}(r)}{r} \end{aligned} \quad (2.56)$$

The Hartree Energy

The Hartree energy is closely related to the result in Eq. (2.56)

$$\begin{aligned} E_H &= \sum_{i,j} \int d\mathbf{r}' d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}') \psi_i(\mathbf{r}) \\ &= \sum_i \int_0^\infty dr u_{n_i l_i}^*(r) \hat{V}_H u_{n_i l_i}(r), \end{aligned} \quad (2.57)$$

where the summation includes the $i = j$ self-interaction term. This equation can be simplified by use of the identity ³

$$R_k(ij; rt) = \int_0^\infty dr u_{n_i l_i}^*(r) u_{n_r l_r}(r) \frac{1}{r} \mathcal{Y}_k(n_j l_j; n_t l_t/r), \quad (2.58)$$

³See Eq. (13-19) on pg. 310 of [16]

to yield the result

$$E_H = \sum_{i,j} \sum_{k=0}^{\min(2l_i, 2l_j)} (-1)^{-m_i - m_j} \frac{(2l_j + 1)(2l_i + 1)}{(2k + 1)^2} \langle l_j m_j l_j (-m_j) | k 0 \rangle \langle l_j 0 l_j 0 | k 0 \rangle \times \\ \langle l_i m_i l_i (-m_i) | k 0 \rangle \langle l_i 0 l_i 0 | k 0 \rangle \mathcal{R}_k(ij; ij) \quad (2.59)$$

It is convenient to rewrite the result for the Hartree energy in terms of the Gaunt coefficients

$$c^k(l_1 m_1; l_2 m_2) = \frac{2l_2 + 1}{2l_1 + 1} \langle k 0 l_2 0 | l_1 0 \rangle \langle k(m_1 - m_2) l_2 m_2 | l_1 m_1 \rangle, \quad (2.60)$$

such that

$$E_H = \sum_{i,j} \sum_{k=0}^{\min(2l_i, 2l_j)} c^k(l_i m_i; l_i m_i) c^k(l_j m_j; l_j m_j) \mathcal{R}_k(ij; ij) \quad (2.61)$$

2.4.4 The Exchange Term

The Exchange operator acting on the i th orbital

$$\hat{V}_X \psi_i(\mathbf{r}) = - \sum_j \delta_{\sigma_i, \sigma_j} \int d\mathbf{r}' \frac{\psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}), \quad (2.62)$$

can be written in terms of the radial functions $u_i(r)$ by following the same steps as outlined in deriving the Hartree term.

Applying the identities in Eqs. (2.42) and (2.43) leads to

$$\hat{V}_X \psi_i(\mathbf{r}) = - \sum_j \delta_{\sigma_i, \sigma_j} \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{4\pi}{2k+1} Y_{kq}(\theta, \phi) \int_0^{\infty} dr' u_{n_j l_j}^*(r') u_{n_i l_i}(r') \frac{r_{<}^k}{r_{>}^{k+1}} \times \\ \left[\int d\Omega' [Y_{l_j, m_j}(\theta', \phi')]^* Y_{l_i, m_i}(\theta', \phi') Y_{kq}(\theta', \phi') \right] \frac{u_{n_j l_j}(r)}{r} Y_{l_j, m_j}(\theta, \phi)$$

Now substitute the radial integral in Eq. (2.47) to yield

$$\begin{aligned} \hat{V}_X \psi_i(\mathbf{r}) = & - \sum_j \delta_{\sigma_i, \sigma_j} \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{4\pi}{2k+1} Y_{kq}(\theta, \phi) Y_{l_j, m_j}(\theta, \phi) \frac{\mathcal{Y}_k(n_j l_j; n_i l_i / r)}{r} \frac{u_{n_j l_j}^*(r)}{r} \times \\ & \sqrt{\frac{(2k+1)(2l_j+1)}{4\pi(2l_i+1)}} \langle l_j m_j k q | l_i m_i \rangle \langle l_j 0 k 0 | l_i 0 \rangle \end{aligned}$$

with the following restrictions

$$\begin{aligned} m_j + q &= m_i & \implies q = m_i - m_j \\ l_j + k + l_i &= \text{even, else } 0 \\ l_j + k &\geq l_i \geq |l_j - k| & \implies |l_j - l_i| \leq k \leq l_j + l_i. \end{aligned} \quad (2.63)$$

To eliminate any dependence on θ or ϕ perform the integration

$$\int d\Omega [Y_{l_i, m_i}(\theta, \phi)]^* \left[- \sum_j \delta_{\sigma_i, \sigma_j} \int d\mathbf{r}' \frac{\psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}) \right] \quad (2.64)$$

which leads to

$$\begin{aligned} = & - \sum_j \delta_{\sigma_i, \sigma_j} \sum_{k=0}^{\infty} \sum_{q=-k}^k \sqrt{\frac{4\pi(2l_j+1)}{(2k+1)(2l_i+1)}} \frac{\mathcal{Y}_k(n_j l_j; n_i l_i / r)}{r} \frac{u_{n_j l_j}(r)}{r} \times \\ & \langle l_j m_j k q | l_i m_i \rangle \langle l_j 0 k 0 | l_i 0 \rangle \int d\Omega (Y_{l_i, m_i}(\theta, \phi))^* Y_k^q(\theta, \phi) Y_{l_j, m_j}(\theta, \phi) \\ = & - \sum_j \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{2l_j+1}{2l_i+1} \langle l_j m_j k q | l_i m_i \rangle \langle l_j 0 k 0 | l_i 0 \rangle \times \\ & \langle k q l_j m_j | l_i m_i \rangle \langle k 0 l_j 0 | l_i 0 \rangle \frac{\mathcal{Y}_k(n_j l_j; n_i l_i / r)}{r} \frac{u_{n_j l_j}(r)}{r} \end{aligned}$$

with the additional restrictions

$$\begin{aligned}
m_j + q &= m_i \implies q = m_i - m_j \\
l_i + k + l_j &= \text{even, else } 0 \\
l_j + k &\geq l_i \geq |l_j - k| \implies |l_j - l_i| \leq k \leq l_j + l_i.
\end{aligned} \tag{2.65}$$

Utilizing the symmetry properties of the Clebsch-Gordan coefficients, specifically Eq. (2.55), and applying the restrictions yields the final result

$$\begin{aligned}
\hat{V}_X \frac{u_{n_i l_i}(r)}{r} &= - \sum_j \delta_{\sigma_i, \sigma_j} \sum_{k=|l_j-l_i|}^{l_i+l_j} \frac{(2l_i+1)(2l_j+1)}{(2k+1)^2} \langle l_i(-m_i)l_j m_j | k(m_i - m_j) \rangle^2 \times \\
&\quad \langle l_i 0 l_j 0 | k 0 \rangle^2 \frac{\mathcal{Y}_k(n_j l_j; n_i l_i / r)}{r} \frac{u_{n_j l_j}(r)}{r}.
\end{aligned} \tag{2.66}$$

The Exchange Energy

The Exchange energy is closely related to the result from Eq. (2.66)

$$\begin{aligned}
E_X &= - \sum_{i,j} \int d\mathbf{r}' d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}) \psi_i(\mathbf{r}') \\
&= - \sum_i \int_0^\infty dr u_{n_i l_i}^*(r) \hat{V}_X u_{n_i l_i}(r),
\end{aligned} \tag{2.67}$$

where the summation includes $i = j$ self-interaction energy. This equation can be simplified by use of the identity from Eq. (2.58) to yield

$$\begin{aligned}
E_X &= - \sum_j \delta_{\sigma_i, \sigma_j} \sum_{k=|l_j-l_i|}^{l_i+l_j} \frac{(2l_i+1)(2l_j+1)}{(2k+1)^2} \langle l_i(-m_i)l_j m_j | k(m_i - m_j) \rangle^2 \times \\
&\quad \langle l_i 0 l_j 0 | k 0 \rangle^2 \mathcal{R}_k(ij; ji)
\end{aligned} \tag{2.68}$$

In terms of the Gaunt coefficients, Eq. (2.60), the exchange energy has the form

$$E_X = - \sum_j \delta_{\sigma_i, \sigma_j} \sum_{k=|l_j-l_i|}^{l_i+l_j} |c^k(l_i m_i; l_j m_j)|^2 \mathcal{R}_k(ij; ji). \quad (2.69)$$

2.4.5 Transforming the Non-Local Exchange Operator to a Local Operator

The non-local exchange potential in the Hartree-Fock equations represents a difficulty for numerical integration. We would like to transform the non-local exchange operator, if at all possible, to yield an effective local operator:

$$\hat{V}_X \implies \hat{V}_X(r). \quad (2.70)$$

The advantage of such a transformation is that the efficient Numerov algorithm can be used to solve the Hartree-Fock equations. Also, such a transformation facilitates the generation of Hartree-Fock pseudopotentials [17].

The general form of the radially non-local exchange term is given by

$$\hat{V}_X u_{n_i l_i}(r) = \sum_j \delta_{\sigma_i, \sigma_j} \int dr' \sum_k u_{n_j l_j}(r) f_{i,j,k}(r, r') u_{n_j l_j}^*(r') u_{n_i l_i}(r'), \quad (2.71)$$

where the index j is summed over all of the occupied orbitals and the non-local $f_{i,j,k}(r, r')$ correspond to terms in the multi-pole expansion of the Coulomb potential. This equation can be modified by multiplying and dividing by $u_{n_i l_i}(r)$ to yield the local operator

$$\hat{V}_X u_{n_i l_i}(r) = \left[\sum_j \delta_{\sigma_i, \sigma_j} \int dr' \sum_k \left(\frac{u_{n_j l_j}(r)}{u_{n_i l_i}(r)} \right) f_{i,j,k}(r, r') u_{n_j l_j}^*(r') \right] u_{n_i l_i}(r). \quad (2.72)$$

This effective local operator is used in place of the non-local operator in the radial equations. Special care must be taken at the nodes of the radial function $u_{n_i l_i}$ to prevent divergences:

typically a ceiling is applied to the ratio $(u_{n_j l_j}(r)/u_{n_i l_i}(r))$.

2.5 The Scalar-Relativistic Approximation

The purpose of this section is to derive and discuss the ideas behind the scalar-relativistic approximation. The derivation begins with the Dirac equation for an atom; the Dirac equation is the relativistic counterpart for the non-relativistic Schrodinger equation. The scalar-relativistic approximation can be incorporated in SCF theories such as HF and DFT. We are not discussing the relativistic counterparts of HF (Dirac-Fock) or DFT in this thesis.

2.5.1 The Dirac Equation

The Dirac equation (in atomic units)

$$i\frac{\partial\Psi}{\partial t} = (\alpha^{-1}\boldsymbol{\alpha} \cdot \mathbf{p} + \beta\alpha^{-2})\Psi = H\Psi, \quad (2.73)$$

is a relativistically covariant equation with positive definite probability density. The quantity Ψ is a four-component single-particle wave function which describes spin-1/2 particles and $\mathbf{p} = -i\nabla$ is the momentum operator. The matrices $\boldsymbol{\alpha}_i$ and β obey the following algebra

$$\begin{aligned} \alpha_i\alpha_k + \alpha_k\alpha_i &= 2\delta_{ik} \\ \alpha_i\beta + \beta\alpha_i &= 0 \end{aligned} \quad (2.74)$$

$$\alpha_i^2 = \beta^2 = 1_4, \quad (2.75)$$

where 1_4 is the four-dimensional identity matrix. We will choose the following representation for the four Dirac matrices:

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix}, \quad (2.76)$$

where σ_k are the 2×2 Pauli spin matrices and the unit entries of β stand for 2×2 unit matrices. The standard representation for the Pauli matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.77)$$

Coupling a Dirac particle to an electromagnetic interaction of the form

$$A_\mu = (V(r), 0, 0, 0) \quad (2.78)$$

and using the *minimal* substitution (*i.e.* $p^\mu \rightarrow p^\mu - \alpha A^\mu$) reduces the Dirac equation, Eq. (2.73), to

$$i \frac{\partial}{\partial t} \Psi = [\alpha^{-1} \boldsymbol{\alpha} \cdot \mathbf{p} + \beta \alpha^{-2} + V(r)] \Psi. \quad (2.79)$$

The quantity $V(r)$ is assumed to be a spherically symmetric, *i.e.* a function of $r = |\mathbf{r}|$, spin-independent potential. We assume for the wave function the stationary state form

$$\Psi(\mathbf{r}, t) = e^{-i\varepsilon t} \begin{pmatrix} \phi(\mathbf{r}) \\ \chi(\mathbf{r}) \end{pmatrix} \quad (2.80)$$

where $\phi(\mathbf{r}), \chi(\mathbf{r}) \in \mathbb{C}^2$ are time-independent two component spinor describing the spatial and spin-1/2 degrees of freedom. For references on the Dirac equation see [113, 114, 115]

2.5.2 Symmetries of the Motion

The Dirac Hamiltonian, Eq. (2.73), has rotational and inversion symmetry. The total angular momentum operator

$$\mathbf{J} = \begin{pmatrix} \mathbf{L} + \frac{1}{2} \boldsymbol{\sigma} & 0 \\ 0 & \mathbf{L} + \frac{1}{2} \boldsymbol{\sigma} \end{pmatrix}, \quad (2.81)$$

and the parity operator

$$\hat{P} = e^{i\phi}\beta, \quad (2.82)$$

commute with the Hamiltonian. The phase $e^{i\phi}$ equals ± 1 if we require $\hat{P}^2 = 1$. The phase is related to the intrinsic parity of the particle or state and will be denoted n_P . Therefore, the eigenfunctions have well defined angular momentum and parity.

2.5.3 Spin-Orbital Angular Momentum States

The two-component angular solutions, which are simultaneous eigenfunctions of the operators J^2 , J_z , and the spin-orbit coupling term $\sim \mathbf{L} \cdot \mathbf{S}$, can be constructed from the spherical harmonics $Y_{lm}(\hat{\mathbf{r}})$ and the spin-1/2 states

$$\chi_{\frac{1}{2}, \frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{\frac{1}{2}, -\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.83)$$

The solutions have the form

$$\begin{aligned} \varphi_{jm}(l, \frac{1}{2}|\hat{\mathbf{r}}) &= \langle l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} | j, m \rangle Y_{l, m - \frac{1}{2}}(\hat{\mathbf{r}}) \chi_{\frac{1}{2}, \frac{1}{2}} + \\ &\quad \langle l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} | j, m \rangle Y_{l, m + \frac{1}{2}}(\hat{\mathbf{r}}) \chi_{\frac{1}{2}, -\frac{1}{2}}, \end{aligned} \quad (2.84)$$

with

$$m' = m \pm \sigma; \quad \sigma = \pm \frac{1}{2}. \quad (2.85)$$

Since the quantum numbers (lm') are integers, (jm) are half-integers. From the triangle inequalities $|l - 1/2| \leq j \leq |l + 1/2|$, we conclude $j = l \pm \frac{1}{2}$, or equivalently $l = j \mp 1/2$. This means that a given value of j corresponds to two different values of l , or a given value of l corresponds to two values of j : the exception being the case of $l = 0$, which has a unique value of $j = 1/2$.

Using the proper Clebsch-Gordan coefficients, the spin-orbital angular momentum states

are: for $j = l + 1/2$

$$\varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) = \sqrt{\frac{l + 1/2 + m}{2l + 1}} Y_{l, m - \frac{1}{2}}(\hat{\mathbf{r}}) \chi_{\frac{1}{2}, +\frac{1}{2}} + \sqrt{\frac{l + 1/2 - m}{2l + 1}} Y_{l, m + \frac{1}{2}}(\hat{\mathbf{r}}) \chi_{\frac{1}{2}, -\frac{1}{2}} \quad (2.86)$$

and for $j = l - 1/2$

$$\varphi_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) = \sqrt{\frac{l + 1/2 - m}{2l + 1}} Y_{l, m - \frac{1}{2}}(\hat{\mathbf{r}}) \chi_{\frac{1}{2}, +\frac{1}{2}} - \sqrt{\frac{l + 1/2 + m}{2l + 1}} Y_{l, m + \frac{1}{2}}(\hat{\mathbf{r}}) \chi_{\frac{1}{2}, -\frac{1}{2}}. \quad (2.87)$$

2.5.4 Eigenvalues of the Spin-Orbital Angular Momentum States

For the states in Eqs. (2.86, 2.87), it holds

$$J^2 \varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) = j(j + 1) \varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}), \quad (2.88)$$

and

$$J_z \varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) = m \varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}). \quad (2.89)$$

Furthermore,

$$J^2 = (\mathbf{L} + \mathbf{S})^2 = L^2 + S^2 + 2\mathbf{L} \cdot \mathbf{S}, \quad (2.90)$$

or equivalently, using the identity $\mathbf{S} = \boldsymbol{\sigma}/2$

$$\mathbf{L} \cdot \boldsymbol{\sigma} = J^2 - L^2 - \frac{3}{4}. \quad (2.91)$$

For $j = l + 1/2$

$$\begin{aligned} \mathbf{L} \cdot \boldsymbol{\sigma} \varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) &= \left[j(j + 1) - (j - \frac{1}{2})(j - \frac{3}{2}) - \frac{3}{4} \right] \varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \\ &= \left[j - \frac{1}{2} \right] \varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}), \end{aligned} \quad (2.92)$$

and for $j = l' - 1/2$

$$\begin{aligned}\mathbf{L} \cdot \boldsymbol{\sigma} \varphi_{jm}(j + \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) &= \left[j(j+1) - (j + \frac{1}{2})(j - \frac{1}{2}) - \frac{3}{4} \right] \varphi_{jm}(j + \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) \\ &= \left[-j - \frac{3}{2} \right] \varphi_{jm}(j + \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}).\end{aligned}\tag{2.93}$$

2.5.5 Parity and the Operator $\sigma \cdot \hat{\mathbf{r}}$

The effect of the parity operator, Eq. (2.82), on the state of the form

$$\Psi(\mathbf{r}, t) = \begin{pmatrix} \phi(\mathbf{r}, t) \\ \chi(\mathbf{r}, t) \end{pmatrix},\tag{2.94}$$

is given by

$$\hat{P}\Psi(\mathbf{r}, t) = n_P \beta \Psi(\mathbf{r}, t) = \Psi'(-\mathbf{r}, t),\tag{2.95}$$

which leads to

$$\Psi'(\mathbf{r}, t) = n_P \begin{pmatrix} \phi(-\mathbf{r}, t) \\ -\chi(-\mathbf{r}, t) \end{pmatrix}.\tag{2.96}$$

For a state with definite parity (*i.e.* it is unchanged by the transformation $\Psi' = \Psi$) we have the condition

$$\begin{aligned}\phi(\mathbf{r}, t) &= n_P \phi(-\mathbf{r}, t) \\ \chi(\mathbf{r}, t) &= -n_P \chi(-\mathbf{r}, t).\end{aligned}\tag{2.97}$$

The upper component of the Dirac wave function has the same spatial parity as that of the overall state, while the lower component has the opposite spatial parity. This is due to the action of the β matrix, which gives an extra phase to the χ spinor. An operator of odd parity, which is a linear combination of $Y_{lm}(\hat{\mathbf{r}})$, can be used to transform one state into the other, since their l values differ by 1.

The inversion properties of the spherical harmonics state that when $\mathbf{r} = (r, \theta, \phi)$ is replaced by $-\mathbf{r} = (r, \pi - \theta, \pi + \phi)$:

$$Y_{lm}(\pi - \theta, \pi + \phi) = (-1)^l Y_{lm}(\theta, \phi). \quad (2.98)$$

For a given j , our spin-orbit angular solutions $\varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}})$, $\varphi_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}})$ are of opposite parity (*i.e.* their property to change only by a factor of ± 1 under spatial inversion) since their l values differ by 1.

The pseudo-scalar $\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}$ has odd parity:

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{r}} \rightarrow \boldsymbol{\sigma} \cdot -\hat{\mathbf{r}} = -\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}. \quad (2.99)$$

In addition $\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}$ is Hermitian and when applied to the states $\varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}})$ changes their parity. It can be shown that in the two-dimensional basis given by $\varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}})$, $\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}$ has the diagonal form

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{r}} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad (2.100)$$

The result can be written as

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{r}} \varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) = -\varphi_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}). \quad (2.101)$$

2.5.6 The Operator $\boldsymbol{\sigma} \cdot \mathbf{p}$

The operator $\boldsymbol{\sigma} \cdot \mathbf{p}$ plays an important role in the radial Dirac equation. We want to determine its action on the functions of the form $a(r)\varphi_{jm}(j \mp \frac{1}{2}|\hat{\mathbf{r}})$. Using $\mathbf{p} = -i\nabla$ and the following identity for the Pauli spin matrices

$$(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} 1_2 + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}). \quad (2.102)$$

It is possible to obtain the relations

$$(\boldsymbol{\sigma} \cdot \hat{\mathbf{r}})^2 = 1_2 \quad (2.103)$$

and

$$\begin{aligned} \boldsymbol{\sigma} \cdot \mathbf{p} &= \frac{(\boldsymbol{\sigma} \cdot \mathbf{r})}{r^2} (\boldsymbol{\sigma} \cdot \mathbf{r} \boldsymbol{\sigma} \cdot \mathbf{p}) \\ &= \frac{(\boldsymbol{\sigma} \cdot \mathbf{r})}{r} \left(\frac{1}{i} \frac{d}{dr} + i \frac{\boldsymbol{\sigma} \cdot \mathbf{L}}{r} \right). \end{aligned} \quad (2.104)$$

We can now write: for $j = l + 1/2$

$$\begin{aligned} \boldsymbol{\sigma} \cdot \mathbf{p} a(r) \varphi_{jm}(j - \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) &= \frac{(\boldsymbol{\sigma} \cdot \mathbf{r})}{r} \left(\frac{1}{i} \frac{d}{dr} + i \frac{(j - \frac{1}{2})}{r} \right) \varphi_{jm}(j - \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) \\ &= -i \left(\frac{da(r)}{dr} - \frac{(j - \frac{1}{2})}{r} a(r) \right) \frac{(\boldsymbol{\sigma} \cdot \mathbf{r})}{r} \varphi_{jm}(j - \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) \end{aligned} \quad (2.105)$$

and for $j = l' - 1/2$

$$\begin{aligned} \boldsymbol{\sigma} \cdot \mathbf{p} a(r) \varphi_{jm}(j + \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) &= \frac{(\boldsymbol{\sigma} \cdot \mathbf{r})}{r} \left(\frac{1}{i} \frac{d}{dr} - i \frac{(j + \frac{3}{2})}{r} \right) \varphi_{jm}(j + \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) \\ &= -i \left(\frac{da(r)}{dr} + \frac{(j + \frac{3}{2})}{r} a(r) \right) \frac{(\boldsymbol{\sigma} \cdot \mathbf{r})}{r} \varphi_{jm}(j + \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) \end{aligned} \quad (2.106)$$

2.5.7 The Radial Dirac Equation

Using the state in Eq. (2.80) will produce the time-independent Dirac equation

$$H\psi = \varepsilon\psi \quad (2.107)$$

where ψ is the two-component spinor

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}. \quad (2.108)$$

In the case of electrons (positive energy solutions) ϕ is the large component and χ is the small component (*i.e* by a factor of α). Solving the eigenvalue problem Eq. (2.108) leads to a set of coupled equation for ϕ and χ :

$$\begin{aligned} \alpha^{-1}(\boldsymbol{\sigma} \cdot \mathbf{p})\chi &= (\varepsilon - V - \alpha^{-2})\phi \\ \alpha^{-1}(\boldsymbol{\sigma} \cdot \mathbf{p})\phi &= (\varepsilon - V + \alpha^{-2})\chi. \end{aligned} \quad (2.109)$$

The coupling is due to the $\boldsymbol{\sigma} \cdot \mathbf{p}$ term, which has odd parity and does not couple states with different sets of (jm) quantum numbers. The general solution can be characterized by its parity $(-1)^l$ and quantum numbers (jm) :

$$\psi_{njm} = \begin{pmatrix} g_{nj}(r)\varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \\ -if_{nj}(r)\frac{\boldsymbol{\sigma} \cdot \mathbf{r}}{r}\varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} g_{nj}(r)\varphi_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \\ -if_{nj}(r)\frac{\boldsymbol{\sigma} \cdot \mathbf{r}}{r}\varphi_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \end{pmatrix}. \quad (2.110)$$

Note that Eq. (2.110) incorporates the results of Eq. (2.97) by explicitly using φ 's with opposite parity to describe the upper and lower components. Using Eqs. (2.101) allows Eq. (2.110) to be simplified to

$$\psi = \begin{pmatrix} g_{nj}(r)\varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \\ if_{nj}(r)\varphi_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} g_{nj}(r)\varphi_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \\ if_{nj}(r)\varphi_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{\mathbf{r}}) \end{pmatrix} \quad (2.111)$$

Applying Eq. (2.111) to Eq. (2.109) leads to: for $l = j - 1/2$ and $l' = j + 1/2$:

$$\begin{aligned}\alpha^{-1} \left(\frac{df_{nj}}{dr} - \frac{(j - \frac{1}{2})}{r} f_{nj} \right) &= -(\varepsilon - V - \alpha^{-2}) g_{nj} \\ \alpha^{-1} \left(\frac{dg_{nj}}{dr} + \frac{(j + \frac{3}{2})}{r} g_{nj} \right) &= (\varepsilon - V + \alpha^{-2}) f_{nj}.\end{aligned}\tag{2.112}$$

and for $l = j + 1/2$ and $l' = j - 1/2$:

$$\begin{aligned}\alpha^{-1} \left(\frac{df_{nj}}{dr} + \frac{(j + \frac{3}{2})}{r} f_{nj} \right) &= -(\varepsilon - V - \alpha^{-2}) g_{nj} \\ \alpha^{-1} \left(\frac{dg_{nj}}{dr} - \frac{(j - \frac{1}{2})}{r} g_{nj} \right) &= (\varepsilon - V + \alpha^{-2}) f_{nj}.\end{aligned}\tag{2.113}$$

At this point it is advantageous to introduce the quantum number κ , which is related to the quantum numbers j, l and l' by

$$\begin{aligned}\text{if } \kappa &= j + \frac{1}{2}, \quad \text{then } l = j + \frac{1}{2} \text{ and } l' = j - \frac{1}{2} \Rightarrow \kappa = l, \\ \text{if } \kappa &= -(j + \frac{1}{2}), \quad \text{then } l = j - \frac{1}{2} \text{ and } l' = j + \frac{1}{2} \Rightarrow \kappa = -l - 1,\end{aligned}\tag{2.114}$$

which allows Eqs. (2.112) and (2.113) to be written concisely as

$$\begin{aligned}\alpha^{-1} \left(\frac{df_{n\kappa}}{dr} - \frac{(\kappa - 1)}{r} f_{n\kappa} \right) &= -(\varepsilon - V - \alpha^{-2}) g_{n\kappa} \\ \alpha^{-1} \left(\frac{dg_{n\kappa}}{dr} + \frac{(\kappa + 1)}{r} g_{n\kappa} \right) &= (\varepsilon - V + \alpha^{-2}) f_{n\kappa}.\end{aligned}\tag{2.115}$$

By defining a new energy origin

$$\varepsilon = \varepsilon - \alpha^{-2},\tag{2.116}$$

the factor multiplying f on the right-hand side can be written as $2M/\alpha^2$, with

$$M(r) = 1 + \frac{\alpha^2}{2} [\varepsilon - V(r)].\tag{2.117}$$

Using Eqs. (2.116) and (2.117) it is now possible to rewrite Eq. (2.115) as

$$\frac{df_{n\kappa}}{dr} = \alpha(V - \varepsilon)g_{n\kappa} + \frac{(\kappa - 1)}{r}f_{n\kappa} \quad (2.118)$$

$$\frac{dg_{n\kappa}}{dr} = -\frac{(\kappa + 1)}{r}g_{n\kappa} + \frac{2M}{\alpha}f_{n\kappa} \quad (2.119)$$

Solving the second equation for f and substituting this into the first equation and carrying out the differentiation, we obtain

$$\begin{aligned} & -\frac{1}{2M} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dg_{n\kappa}}{dr} \right) + \left[V + \frac{1}{2M} \frac{\kappa(\kappa + 1)}{r^2} \right] g_{n\kappa} - \frac{\alpha^2}{4M^2} \frac{dV}{dr} \frac{dg_{n\kappa}}{dr} \\ & - \frac{\alpha^2}{4M^2} \frac{dV}{dr} \frac{(1 + \kappa)}{r} g_{n\kappa} = \varepsilon g_{n\kappa}. \end{aligned} \quad (2.120)$$

For each of the cases in Eq. (2.114), the factor of $\kappa(\kappa + 1)$ can be written as

$$\kappa(\kappa + 1) = l(l + 1) \quad (2.121)$$

Apart from the last two terms on the left-hand side, Eq. (2.120) looks like a radial Schrodinger equation. No approximations have been made yet. The formal mass term M is sometimes called the *mass-velocity* term, the term $(dV/dr)(dg/dr)$ is known as the *Darwin* term, and the last term on the left-hand side of Eq. (2.120) is the *spin-orbit coupling* term; this is true because

$$\mathbf{L} \cdot \boldsymbol{\sigma} \varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}) = -(1 + \kappa) \varphi_{jm}(j \mp \frac{1}{2}, \frac{1}{2} | \hat{\mathbf{r}}). \quad (2.122)$$

2.5.8 The Koelling-Harmon Approximation

The essential approximation (Koelling-Harmon) [116] consists of dropping the spin-orbit terms from the radial equations, Eq. (2.120). The spin-orbit interaction is separated out explicitly and is normally treated variationally or by perturbation theory. The approximate

functions are called \tilde{g} and \tilde{f} , and take the form

$$-\frac{1}{2M} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\tilde{g}_{nl}}{dr} \right) + \left[V + \frac{1}{2M} \frac{l(l+1)}{r^2} \right] \tilde{g}_{nl} - \frac{\alpha^2}{4M^2} \frac{dV}{dr} \frac{d\tilde{g}_{nl}}{dr} = \varepsilon \tilde{g}_{nl} \quad (2.123)$$

and

$$\tilde{f}_{nl} = \frac{\alpha}{2M} \frac{d\tilde{g}_{nl}}{dr}. \quad (2.124)$$

The latter is needed for normalization

$$\int dr (\tilde{g}_{nl}^2 + \tilde{f}_{nl}^2) r^2 = 1 \quad (2.125)$$

It should be noted that some versions of the scalar-relativistic approximation ignore the small component \tilde{f}_{nl} [116, 17]. Eq. (2.123) is the scalar-relativistic radial equation, and it may be used in place of the Schrodinger equation. Note that the $(n\kappa)$ index has been replaced by (nl) since there is no longer any j dependence.

The spin-orbit coupling term

$$H_{SO} = \frac{\alpha^2}{4M^2} \frac{1}{r} \frac{dV}{dr} \mathbf{L} \cdot \boldsymbol{\sigma} \quad (2.126)$$

can be treated as a perturbation: the factor $\frac{1}{r} \frac{dV}{dr}$ in Eq. (2.126) is largest near the nucleus. The evaluation of the spin-orbit term with respect to the scalar-relativistic orbitals is discussed in Ref. [116].

2.5.9 Solving the Radial Equations

We would like to transform Eq. (2.123) to a form that can be solved by the Numerov algorithm, see Appendix A.5. This is possible by the change of variables: let $u(r) = r\tilde{g}(r)$,

with the derivatives $u' = \tilde{g} + r\tilde{g}'$ and $u'' = 2\tilde{g}' + r\tilde{g}''$, which yields

$$-\frac{1}{2M} \frac{d^2 u}{dr^2} + \left[V + \frac{1}{2M} \frac{l(l+1)}{r^2} + \frac{\alpha^2}{4M^2 r} \frac{dV}{dr} \right] u - \frac{\alpha^2}{4M^2} \frac{dV}{dr} \frac{du}{dr} = \varepsilon u \quad (2.127)$$

Now let $Z(r) = u(r)/\sqrt{M}$, this leads to

$$\frac{d^2 Z}{dr^2} = \frac{1}{M^{1/2}} \frac{d^2 u}{dr^2} + \frac{\alpha^2}{2M^{3/2}} \frac{dV}{dr} \frac{du}{dr} + \frac{3\alpha^4}{16M^{5/2}} \left(\frac{dV}{dr} \right)^2 u + \frac{\alpha^2}{4M^{3/2}} \frac{d^2 V}{dr^2} u, \quad (2.128)$$

and finally

$$\frac{d^2 Z}{dr^2} = \left[\frac{l(l+1)}{r^2} + 2M(V - \varepsilon) + \frac{\alpha^2}{2Mr} \frac{dV}{dr} + \frac{3\alpha^4}{16M^2} \left(\frac{dV}{dr} \right)^2 + \frac{\alpha^2}{4M} \frac{d^2 V}{dr^2} \right] Z. \quad (2.129)$$

2.6 Results for Germanium

This section presents the results from our numerical Hartree-Fock code for atomic Germanium. Figure 7.1(a) plots the radial orbitals for Ge (3P) ground-state, and Figure 7.1(b) shows a close up of the valence orbitals. Table 2.1 compares the eigenvalues for non-relativistic and scalar-relativistic HF.

2.7 Time-Dependent Density Function Theory

Time-Dependent Density Function Theory (TDDFT) is a generalization of the static Hohenberg-Kohn-Sham theory to the time domain. In principle TDDFT can describe the excitation spectra exactly [117]; however, TDDFT is plagued by the same problem as static DFT via the exchange-correlation functional. The exact functional is not known and the accuracy of practical approximations is not established [118].

For references on TDDFT see Refs. [117, 118, 29, 119, 120, 121]. In static DFT, the ground-state of a system is determined by the minimum of the total energy. For the case

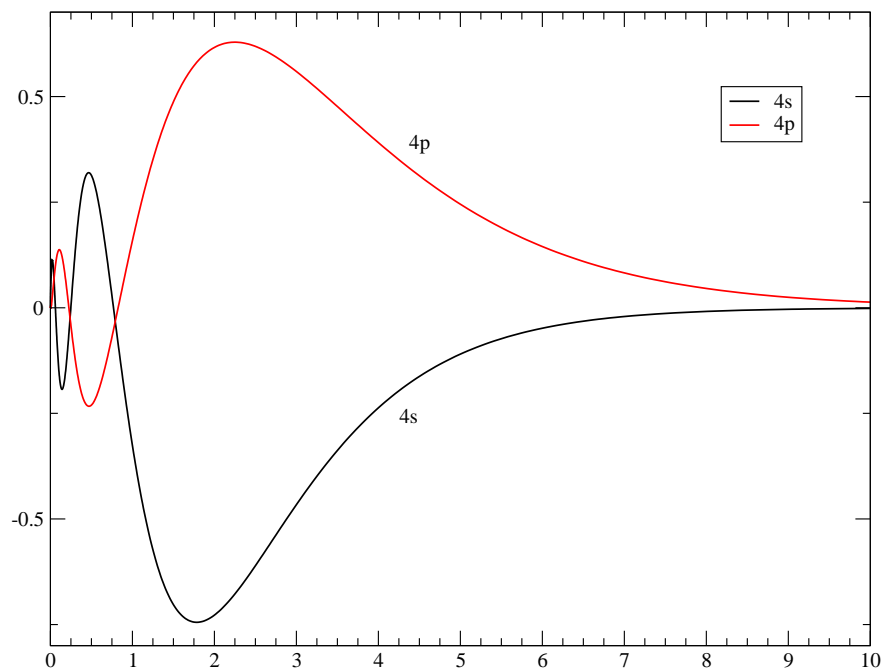
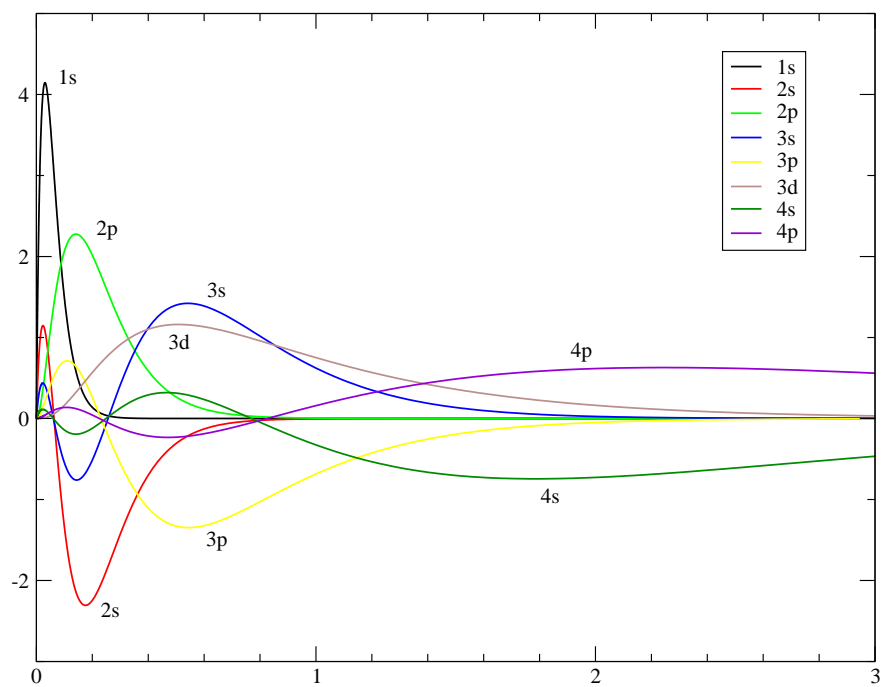


Figure 2.1: 7.1(a) The radial orbitals for the (3P) ground-state of the Ge atom from all-electron scalar-relativistic Hartree-Fock. 7.1(b) The valence 4s and 4p orbitals.

Table 2.1: Comparison of the results of our radial HF code for scalar-relativistic (SR) and non-relativistic HF (NR) for the ground-state of the Ge atom (3P). We also compare our non-relativistic results to the reference data of Fischer [4].

(a) Comparison of the eigenvalues.

shell	SR	NR	NR Ref. [4]
1s	-411.18370	-405.24423	-405.24445
2s	-53.46802	-52.15020	-52.15034
2p	-46.56265	-46.23602	-46.23616
3s	-7.40791	-7.19093	-7.19100
3p	-5.21133	-5.16153	-5.16160
3d	-1.59891	-1.63483	-1.63490
4s	-0.56670	-0.55336	-0.55336
4p	-0.28614	-0.28735	-0.28735

(b) Comparison of the total energy.

	SR	NR	NR Ref. [4]
Total Energy	-2097.4798	-2075.3581	-2075.3597

of a system under the influence of a time-dependent external-potential, the evolution of the system is determined by the extremum of the quantum-mechanical action:

$$A = \int_{t_0}^{t_1} dt \langle \psi(t) | \left[i \frac{d}{dt} - \hat{H}(t) \right] | \psi(t) \rangle. \quad (2.130)$$

This is the stationary action principle and is one of the essential theorems of TDDFT by Runge and Gross [117]. The second theorem is that the time-dependent density determines the external potential uniquely up to an additive time-dependent function. A consequence of this one-to-one mapping between the time-dependent potential and the time-dependent density is that the expectation value of any quantum mechanical operator is a unique functional of the density. In the work of Runge and Gross the external potential must be such that it can be expanded as a Taylor series in time [117]. An interesting feature of TDDFT is that the density functionals depend on the initial state of the system.

The quantum mechanical action is a functional of the density, $A[n]$, and has a stationary point at the correct time-dependent density. In analogy to the Kohn-Sham approach, it

is possible to introduce an auxiliary non-interacting system with the same time-dependent density as the interacting system $n(\mathbf{r}, t)$. The form of the action functional is given by

$$A[n] = \sum_i \int_{t_0}^{t_1} dt \langle \psi_i(t) | \left[i \frac{d}{dt} + \frac{1}{2} \nabla^2 - V_{ext}(t) \right] | \psi_i(t) \rangle - A_H[n] - A_{xc}[n], \quad (2.131)$$

where A_H and A_{xc} are the time-dependent Hartree and exchange-correlation terms. Following a similar approach used in ground-state DFT, it is possible to derive the following set of time-dependent Kohn-Sham equations

$$H(t)\psi_i(t) = i \frac{d\psi_i}{dt}, \quad H(t) = -\frac{1}{2} \nabla^2 + V_{eff}(\mathbf{r}, t), \quad (2.132)$$

with the density given by

$$n(\mathbf{r}, t) = \sum_i |\psi_i(\mathbf{r}, t)|^2 \quad (2.133)$$

and the initial state

$$\Psi(t_0) = \Psi_0. \quad (2.134)$$

The effective time-dependent potential felt by the electrons is

$$V_{eff} = -\frac{1}{2} \nabla^2 + V_{ext}(\mathbf{r}, t) + \int d\mathbf{r}' \frac{n(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} + V_{xc}[n](\mathbf{r}, t), \quad (2.135)$$

where the time-dependent exchange-correlation functional is the variational derivative of the exchange-correlation part of the action functional

$$V_{xc}[n](\mathbf{r}, t) = \frac{\delta A_{xc}[n]}{\delta n(\mathbf{r}, t)}. \quad (2.136)$$

Note that $V_{xc}[n](\mathbf{r}, t)$ is a function of \mathbf{r} and t and a functional of the time-dependent density $n(\mathbf{r}', t')$ [29].

2.7.1 The Adiabatic Local Density Approximation

The time-dependent exchange-correlation functional $V_{xc}[n](\mathbf{r}, t)$ is of course unknown. The most commonly used approximation, largely due to its simplicity, is the Adiabatic Local Density Approximation (ALDA). In the ALDA the exchange-correlation functional is the static LDA exchange-correlation functional; there is no time (frequency) dependence. This can be justified for low frequencies or slowly varying densities [122, 29, 118, 123].

2.7.2 Linear Response

TDDFT can be used to calculate neutral excitations of a system by finding the poles of the linear density response function. The *interacting* linear density response function

$$\chi = \left. \frac{\delta n(\mathbf{r}, t)}{\delta V_{ext}(\mathbf{r}', t')} \right|_{V_{ext}=0}, \quad (2.137)$$

measures the degree to which the density responds to first order in the external potential. It is possible to derive a Dyson like equation for χ in terms of the *non-interacting* Kohn-Sham response

$$\chi_0 = \left. \frac{\delta n(\mathbf{r}, t)}{\delta V_{eff}(\mathbf{r}', t')} \right|_{V_{eff}=0}, \quad (2.138)$$

where $V_{eff}[n] = V_{ext} + V_H[n] + V_{xc}[n]$. The linear response of the fictitious Kohn-Sham system has poles at the Kohn-Sham eigenvalue differences. The relationship between χ and χ_0 is given by:

$$\chi(\mathbf{r}, \mathbf{r}', \omega) = \chi_0(\mathbf{r}, \mathbf{r}', \omega) + \int d\mathbf{r}_1 d\mathbf{r}_2 \chi_0(\mathbf{r}, \mathbf{r}_1, \omega) K(\mathbf{r}_1, \mathbf{r}_2, \omega) \chi(\mathbf{r}_2, \mathbf{r}', \omega) \quad (2.139)$$

where the interaction kernel is

$$K(\mathbf{r}_1, \mathbf{r}_2, \omega) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + \left. \frac{\delta V_{xc}[n(\mathbf{r}, t)]}{\delta n(\mathbf{r}', t')} \right|_{V_{ext}=0}. \quad (2.140)$$

The first term in the interaction kernel is the Coulomb contribution and the second term is the time-dependent exchange-correlation contribution. In the ALDA the time-dependent exchange-correlation kernel $\delta V_{xc}[n(\mathbf{r}, t)] \delta n(\mathbf{r}', t')$ is local in both time and space.

Eq. (2.139) can be cast in the form of an eigenvalue equation involving pairs of occupied and unoccupied Kohn-Sham single-particle states and their respective eigenvalues.

2.7.3 Explicit Real Time Calculations

It is possible to carry out real-time calculations by explicitly evolving the single-particle orbitals $\{\psi_i\}$ [1, 119, 123, 124, 29, 125]. This approach has the advantage over the linear response approach (see previous section) in that it is not limited to small perturbations and can be used for non-linear effects. In addition, the real-time approach has an advantage for larger systems since only the occupied states are propagated.

From the real-time approach it is possible to calculate the polarizability $\alpha(\omega)$, which is one of the most important response functions because it describes the distortion of a charge cloud by an external field. The polarizability is related to the dipole strength function $S(\omega)$:

$$S(\omega) = \frac{2m}{\pi e^2 \hbar} \omega \Im \alpha(\omega). \quad (2.141)$$

The dipole strength function $S(\omega)$ is a useful quantity because it is proportional to the experimentally measured photoabsorption cross section $\sigma(\omega)$. The relationship between S and σ is given by (in SI units):

$$S(\omega) = \frac{2m}{\pi e^2 \hbar} \omega \Im \alpha(\omega). \quad (2.142)$$

In addition, integration of S with respect to the energy gives the number of electrons:

$$\int_0^\infty dE S(E) = \sum_i f_i = N, \quad (2.143)$$

where the f_i are the oscillator strengths.

A way to calculate $\alpha(\omega)$ is outlined in Refs. [123, 1, 119, 29]. For time $t < 0$, the Hamiltonian H includes a perturbation due to an electric field, $\Delta H = \mathbf{E} \cdot \mathbf{x}$, along the x -direction. Normal time-independent DFT is used to solve the Kohn-Sham equations for the ground-state of this system. At time $t = 0$, the electric field is switched off and the occupied Kohn-Sham eigenstates are evolved in time by solving Eq. (2.132) for the set of discrete time steps $t_i = i\Delta t$. For each time step, the electron density is generated by the new set of orbitals, which are then used to calculate the effective Hamiltonian for the next time step. The effective Hamiltonian is a function of time since the density $n(\mathbf{r}, t)$ is a function of time, but for $t > 0$ there is no external time dependence.

For every time step the dipole moment $D(t)$ is calculated:

$$D(t) = \int d\mathbf{r} n(\mathbf{r}, t) = \sum_i \langle \psi_i(t) | x | \psi_i(t) \rangle. \quad (2.144)$$

The Fourier transform of $D(t)$ is given by

$$D(\omega) = \int dt e^{i\omega t - \delta t} D(t), \quad (2.145)$$

where $e^{i\delta t}$ is a damping factor for convergence at large times. To linear order the polarizability is given by

$$D(\omega) = \alpha(\omega) E(\omega) \quad (2.146)$$

The Fourier transform of $E(t) = \overline{E}\Theta(-t)$, where \overline{E} is a constant, is given by

$$E(\omega) = \frac{\overline{E}}{i\omega} \quad (2.147)$$

which yields the relation

$$\Im \alpha(\omega) = \frac{\omega \Re D(\omega)}{\overline{E}}. \quad (2.148)$$

2.8 Broken Symmetry Solutions

Typically the mean-field equations occurring in Hartree-Fock or Density Functional Theory are solved in a restricted manner. This means that the spin-orbitals are constrained to have the same spatial function for the spin-up and spin-down spin functions. Due to the exchange interaction between like-spin electrons, each spin-species experiences a different effective potential and would *not* prefer to be described by the same spatial function. For high densities, *i.e.* small values of the Wigner-Seitz parameter r_s , the kinetic energy dominates the electronic structure and restricted solutions typically agree well with those found by an unrestricted approach, where the spin-orbitals have different spatial orbitals for the different spin functions. For lower densities, where the electron-electron interactions dominate, unrestricted solutions produce much better ground state energies reflecting a gain in correlation. Unfortunately the unrestricted approach has the major drawback of producing broken symmetry solutions which are no longer eigenstates of the original Hamiltonian.

An excellent example of this phenomenon is the disassociation problem for the ground-state of the Hydrogen molecule [126]. The Hydrogen molecule disassociates into two non-interacting Hydrogen atoms with arbitrary spin orientation. The energy difference between the singlet ($S=0$) and triplet ($S=1$) states, $\Delta E = E_s - E_t$, must approach zero from below as the molecule disassociates - the singlet always being the ground state. In the restricted Molecular Orbital (MO) approach, the spin-up electron occupies the same molecular orbital as the spin-down electron. These orbitals are used to construct a single determinant wave function which is an eigenstate of the total spin S^2 and z -component S_z , where $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$. Due to this constraint the restricted MO approach fails to yield the correct (H + H) product at large nuclear separation. Using an unrestricted MO approach results in much better energies but produces broken symmetry solutions that are no longer eigenstates of the spin.

The unrestricted HF wave function has the generalized form

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} u(\mathbf{r}_1)\alpha(1) & v(\mathbf{r}_1)\beta(1) \\ u(\mathbf{r}_2)\alpha(2) & v(\mathbf{r}_2)\beta(2) \end{vmatrix}, \quad (2.149)$$

where $u(\mathbf{r})$ and $v(\mathbf{r})$ are the unrestricted spin-up and spin-down orbitals, respectively. By taking a linear combination of Slater determinants of unrestricted MOs, it is possible to restore the symmetry and produce a solution that is a spin eigenstate. The resultant expression for the singlet

$$\Psi_s(\mathbf{r}_1, \mathbf{r}_2) \propto \begin{vmatrix} u(\mathbf{r}_1)\alpha(1) & v(\mathbf{r}_1)\beta(1) \\ u(\mathbf{r}_2)\alpha(2) & v(\mathbf{r}_2)\beta(2) \end{vmatrix} - \begin{vmatrix} u(\mathbf{r}_1)\beta(1) & v(\mathbf{r}_1)\alpha(1) \\ u(\mathbf{r}_2)\beta(2) & v(\mathbf{r}_2)\alpha(2) \end{vmatrix} \quad (2.150)$$

can be reduced to,

$$\Psi_s(\mathbf{r}_1, \mathbf{r}_2) \propto (u(\mathbf{r}_1)v(\mathbf{r}_2) + u(\mathbf{r}_2)v(\mathbf{r}_1))\chi(0,0), \quad (2.151)$$

where $\chi(0,0) = (\alpha(1)\beta(2) - \alpha(2)\beta(1))/\sqrt{2}$ is the spin eigenfunction for the singlet [127, 128, 129]. This solution has the form of a Heitler-London [130] or valance bond wave function, which can be shown to be of lower energy than the original unrestricted solution.

The important point is that the orbitals u and v are solutions to the mean-field equation and *not* the original many-body Hamiltonian. For this reason mean-field methods give sharp transitions, whereas a correct many-body treatment gives a continuous variation. It is well known in nuclear physics [131] that a *simple* wave function, *i.e.* a product of single-particle orbitals, cannot properly describe correlations between the electrons if the proper symmetry behavior is simultaneously required.

Chapter 3

Theoretical Methods: Many-Body Methods

3.1 Introduction

It is well known that single-body methods such as Density Functional Theory (DFT) in the widely used Local Density Approximation (LDA) underestimates the gaps for excitations, whereas Hartree-Fock (HF) overestimates the gaps [9, 37, 132]. In principle time-dependent DFT (TDDFT) can describe the excitation spectra exactly [117]; however, the exact functional is not known and the accuracy of practical approximations is not established [118]. Thus it is desirable to go beyond single-body theories to a many-body theory.

In this chapter the many-body methods of Configuration Interaction (CI), GW and GW combined with a solution to the Bethe-Salpeter Equation (GW-BSE) are reviewed. The Quantum Monte Carlo method is discussed separately in the succeeding chapter, see chapter 4.

3.2 Configuration Interaction

Exact-diagonalization procedures such as Configuration Interaction (CI) [133] use a linear combination of N -electron Slater determinants to span the Hilbert space of the N -electron system. The form of this expansion is given by

$$\Psi = \sum_i c_i D_i, \tag{3.1}$$

where the D_i are the Slater determinants. In principle CI provides an exact answer for not only the ground state, but also all of the excited states in the limit of a complete basis.

CI is an improvement over the Hartree-Fock method, which incorporates exchange by using an antisymmetric Slater determinant, but neglects the electron correlations resulting from the Coulomb interaction. In practice the number of N -electron determinants is enormous, usually only single and double excitations are included, and one must restrict the basis. Another drawback of this method is that a large number of determinants are needed to describe the gradient discontinuities at the cusps where two electrons have the same position.

A common way to create the D_i are to use the results of an SCF calculation such as Hartree-Fock. An “excited” Slater determinant is created by promoting an electron from an occupied orbital to a virtual (excited) orbital. Below we provide a brief summary of some of the more common CI methods:

- Full CI: The determinants in Eq. (3.1) contain all possible excitations. The number of excitations is determined by the size of the basis.
- Singles-Doubles CI (SDCI): Only single and double excitations are allowed. Singles only CI is also known as CIS.
- Multi-Configuration SCF (MRSCF): Both the CI and SCF coefficients are optimized.
- Complete Active Space SCF (CASSCF): This is a special case of MRSCF, where a set of electrons are designated as being *active*; only excitations involving the active space of electrons are permitted.

Closely related to CI is the coupled-cluster perturbation technique. A common implementation is the coupled-cluster singles-doubles (CCSD) method, which only considers single and double excitations.

3.3 The GW Approximation

GW [37, 134, 9, 118, 17, 135, 39, 38] is a many-body method to describe quasiparticle excitations. A quasiparticle is the interaction of a bare particle, in this case an electron, and a positive *screening* charge surrounding the particle. For electrons the positive screening charge is due to the Coulomb potential; the repulsion between the electrons leads to a depletion of negative charge around a given electron. The mathematical description of quasiparticles is by the single-electron Green's function. More specifically, the quasiparticle energies are the poles of the one-electron Green's function. These energies correspond to addition and removal energies in atoms and molecules, and can be used to compute the band gaps in solids.

GW is a more direct method of calculated energy differences than performing separate total energy calculations. This is because total energy differences are always less accurate than the constituent total energies. This is especially true in QMC method.

GW is directly relevant to this work since the best known calculations of band gaps in Ge and related semiconductors were performed by Shirley, Zhu and Louie [9] using the GW approximation. These calculations were performed with a pseudopotential using the CPP approach in Section 7.7. Also, the CPPs are fit to GW *self-energies* for the lowest bound valence electron of each angular momentum channel of interest outside a bare core [43].

The exact single-electron Green's function is given by

$$G(\mathbf{x}, \mathbf{x}') = \frac{1}{i} \langle N; 0 | T[\Phi(\mathbf{x})\Phi^\dagger(\mathbf{x}')] | N; 0 \rangle, \quad (3.2)$$

where $|N, 0\rangle$ is the interacting ground state for the N electron system, $T[...]$ is the time-ordering operator, $\Phi(\mathbf{x}) = e^{iHt}\Phi(\mathbf{r})e^{-iHt}$ is the fermion annihilation operator in the Heisenberg representation and $\Phi^\dagger(\mathbf{x})$ is the corresponding fermion creation operator [136]. Note that in Eq. (3.2) the coordinate \mathbf{x} represents the spin and space variables: $\mathbf{x} = (\mathbf{r}, t)$. For $t > t'$ the Green's function describes the propagation of a particle added to the many-body

system, while for $t < t'$ the Green's function describes the propagation of a hole added to the many-body system. By the insertion of unity in the given Fock space of the $N + 1$ and $N - 1$ systems ($\sum_i |N + 1; i\rangle\langle N + 1; i|$ and $\sum_j |N - 1; j\rangle\langle N - 1; j|$), the Green's function can be written as

$$G = -i \sum_i \langle N; 0 | \Phi(\mathbf{x}) | N + 1; i \rangle \langle N + 1; i | \Phi^\dagger(\mathbf{x}') | N; 0 \rangle \Theta(t - t') \\ + i \sum_j \langle N; 0 | \Phi^\dagger(\mathbf{x}) | N - 1; j \rangle \langle N - 1; j | \Phi(\mathbf{x}') | N; 0 \rangle \Theta(t' - t), \quad (3.3)$$

where $|N \pm 1; k\rangle$ denotes the k th state of the $N \pm 1$ particle system. Introducing the quasiparticle wave functions and their corresponding energies

$$\phi_i(\mathbf{r}) \langle N; 0 | \Phi(\mathbf{r}) | N + 1; i \rangle, \quad E_i = E_{N+1,i} - E_{N,0} > \mu, \\ \phi_j(\mathbf{r}) \langle N - 1; j | \Phi^\dagger(\mathbf{r}) | N; 0 \rangle, \quad E_i = E_{N,0} - E_{N-1,i} < \mu, \quad (3.4)$$

where μ is the chemical potential, the Green's function becomes

$$G = -i \sum_i \phi_i(\mathbf{r}) \phi_i(\mathbf{r}') \exp[iE_i(t - t')] \Theta(t - t') \\ + i \sum_j \phi_j(\mathbf{r}) \phi_j(\mathbf{r}') \exp[-iE_j(t' - t)] \Theta(t' - t) \quad (3.5)$$

The Green's function is given in the frequency (energy) domain a Fourier transform is with respect to $(t - t')$:

$$G(\mathbf{r}, \mathbf{r}'; E) = \sum_n \frac{\Phi_n(\mathbf{r}) \Phi_n^*(\mathbf{r}')}{E - (E_n \pm i\eta)}, \quad (3.6)$$

where the sum n is over all $(N - 1)$ and $(N + 1)$ particle states and the $\pm i\eta$ terms are for convergence. In this representation the quasiparticle energies correspond to poles of the Green's function.

It is important to note that the quasiparticle energies have an actual physical interpretation, unlike the single-particle eigenvalues from Kohn-Sham DFT which are mathematical

constructs, *i.e.* Lagrange multipliers. The quasiparticles also have finite lifetimes since they are not the exact eigenvalues of the many-body Hamiltonian H . The quasiparticle energies are often regarded as corrections to the Hartree-Fock eigenvalues which approximately describe electron addition and removal energies by Koopmans' theorem, see Section 2.3.2.

The Green's function is a solution to Dyson's equation. A particular form of Dyson's equation (without the spin degrees of freedom)

$$\left[-\frac{1}{2}\nabla^2 + V_H + V_{ext} \right] \Phi_i(\mathbf{r}) + \int d\mathbf{r}' \Sigma(\mathbf{r}, \mathbf{r}'; E_i) \Phi_i(\mathbf{r}') = E_i \Phi_i(\mathbf{r}), \quad (3.7)$$

where V_{ext} is the external potential, in the case of atoms $V_{ext}(r) = -Z/r$, V_H is the Hartree potential

$$V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (3.8)$$

and Σ is the self-energy. The self-energy Σ is a non-Hermitian, energy dependent, non-local operator that describes exchange and correlation beyond the Hartree (V_H) interaction level. In the GW approximation (GWA), Σ is approximated in terms of the Green's function G and the screened interaction W ; in the frequency domain Σ is given by

$$\Sigma(\mathbf{r}, \mathbf{r}'; E) \approx i \int \frac{d\omega}{2\pi} e^{i\eta\omega} G(\mathbf{r}, \mathbf{r}'; E + \omega) W(\mathbf{r}, \mathbf{r}'; -\omega). \quad (3.9)$$

The approximation is the first-order (in W) truncation of an expansion of Hedin's exact coupled many-body equations [135].

3.4 GW-BSE

The GW-BSE method refers to the GW approximation, see Section 3.3, combined with a solution to the Bethe-Salpeter Equation (BSE). This method describes the how a system reacts to the introduction of two interacting quasiparticles (quasielectron and quasihole)

into a system such that bound excitons exist. Therefore accurate promotion energies are accessible.

GW-BSE involves solving an equation that appears to be similar to the Dyson equation in Section 2.7.2 for TDLDA within linear response. The main difference between TDLDA and GW-BSE is the choice of the interaction kernel K and the single-particle states that enter the equations. In TDLDA the single-particle states are the DFT-LDA Kohn-Sham single-particle states, while in GW-BSE the single-particle states are quasiparticle states. The quasiparticle states are obtained from GW and include self-energy effects. Therefore GW-BSE determines excitations of the N -electron system by making reference to the $(N + 1)$ and $(N - 1)$ system, while TDLDA deals with the N -electron system. The BSE interaction kernel couples the quasielectron and quasihole states to describe the exciton [56, 25, 118, 57].

The GW-BSE perturbation approach consists of three successive steps. The first step is performing a DFT-LDA calculation to obtain the single-particle Kohn-Sham states. As a second step the LDA results are used to obtain the self-energy Σ within the GW approximation, which yields the quasiparticle states. Finally, the quasielectron and quasihole states and their respective energies are inserted in the BSE to obtain the optical absorption spectrum.

3.5 Quantum Monte Carlo

For a discussion of Quantum Monte Carlo (QMC) see Chapter 4. For a discussion of many of the technical details of QMC, including the form of the many-body trial wave function, see Chapter 6.

Chapter 4

Quantum Monte Carlo Methods

4.1 Introduction to Monte Carlo

Monte Carlo refers to a broad class of methods use that use random numbers to solve physical problems [137, 7, 106]. One of the most important applications of Monte Carlo is to solve multi-dimensional integrals of the form

$$I = \int d\mathbf{R} f(\mathbf{R}), \quad (4.1)$$

where f is assumed to be a smooth function over the interval and \mathbf{R} is a d -dimensional coordinate. In *quantum Monte Carlo* we are interested in computing integrals of the form

$$\langle \mathcal{A} \rangle = \frac{\int d\mathbf{R} \Pi(\mathbf{R}) \mathcal{A}(\mathbf{R})}{\int d\mathbf{R} \Pi(\mathbf{R})}, \quad (4.2)$$

where Π is a probability distribution and \mathcal{A} is a quantum mechanical operator local in position space. The quantity $\langle \mathcal{A} \rangle$ is a quantum mechanical expectation value of the operator \mathcal{A} with respect to the probability distribution Π .

Typical grid based integration (quadrature) methods evaluate the integral in Eq. (4.1) over a set of N points $\{\mathbf{R}_i; i = 1, \dots, N\}$ with the associated weights w_i :

$$I_N = \frac{1}{N} \sum_{i=1}^N w_i f(\mathbf{R}_i). \quad (4.3)$$

We will assume that the grid is uniform and that the grid-spacing is h . The weights are

independent of the integrand f and are determined by a polynomial expansion; the accuracy of the expansion determines accuracy of the integration method. In general, the error of the integration method can be written in terms of the grid-spacing h as $\mathcal{O}[h^k]$, where $k > 1$. The relationship between the grid-spacing and the number of grid-points N for a d -dimensional hypercube with side L is $N = (L/h)^d$. Therefore, the error of standard quadrature methods scales with respect to the number of integration points as $\mathcal{O}[N^{-k/d}]$.

Monte Carlo methods produce an error that scales (converges) as $\mathcal{O}[N^{-1/2}]$, independent of the dimension of the integral, while using a set of random *configurations* $\{\mathbf{R}_i\}$ with uniform weights $w_i = 1$. In the language of Monte Carlo a configuration or a *walker* represents the set of variables that describes the state of the system: for a multi-dimensional integral the state corresponds to \mathbf{R} , the d -dimensional coordinate. The process by which the walkers are generated is known as a random walk which will be discussed in more detail later.

It quickly becomes clear that Monte Carlo has an advantage over quadrature methods for higher dimensional integrals. In order to reduce the error of a Monte Carlo integration by a factor of two, it is necessary to quadruple the number of walkers (points); this scaling $\propto N^{-1/2}$ is far superior than the scaling standard quadrature methods for high dimensional integrals $\propto N^{-k/d}$. In addition, the number of grid points required to evaluate a multi-dimensional integral can be prohibitively large. For example, to evaluate a 10-dimensional integral with only 10 points per side results in a grid with $N = 10^{10}$ points.

It is imperative to have good random number generators for Monte Carlo simulations. There is no such thing as a truly random number generator, instead we use *pseudo* random number generators. Pseudo random number generators are deterministic and have a finite, although very large, period. All of the work in this thesis uses pseudo random number generators provided by the BOOST [138] and SPRNG [139] libraries.

4.2 The Central Limit Theorem

The Central Limit Theorem (CLT) [140] is the reason why the error of Monte Carlo integration scales as $\mathcal{O}[N^{-1/2}]$, independent of the dimension of the integral, where N is the number of independent configurations used to evaluate (sample) the integral.

In order to understand the CLT, start with a set of uncorrelated configurations $\{\mathbf{R}_i : i = 1, \dots, N\}$ distributed according to a probability distribution function $\mathcal{P}(\mathbf{R})$. Define a new random variable

$$Z_N = \frac{1}{N} \sum_{i=1}^N Z(\mathbf{R}_i), \quad (4.4)$$

where Z is a function of mean

$$\mu_Z = \int d\mathbf{R} \mathcal{P}(\mathbf{R}) Z(\mathbf{R}) \quad (4.5)$$

and variance

$$\sigma_Z^2 = \int d\mathbf{R} \mathcal{P}(\mathbf{R}) [Z(\mathbf{R}) - \mu_Z]^2. \quad (4.6)$$

The variance measures the fluctuations about the mean. In the limit of an infinite number of configurations

$$\mu_Z = \lim_{N \rightarrow \infty} Z_N, \quad (4.7)$$

where μ_Z is the exact mean defined in Eq. (4.5): in practice we would like to estimate the error of $|\mu_Z - Z_N|$.

The CLT states that for large enough N , Z_N is normally distributed about the mean μ_Z with the variance σ_Z^2/N , where σ_Z^2 is the variance of the function Z . The *variance of the mean* is the variance of Z_N , and not of the underlying probability distribution function from which the $\{\mathbf{R}\}$ are drawn. The variance of the mean provides a measure of certainty with which Z_N is equal to μ_Z [31, 7].

In practice the mean μ_Z is measured, as is σ_Z^2 ; therefore the denominator is reduced by

one, leading to the standard formula for the variance of the mean

$$\sigma_{\mu_Z}^2 = \frac{\sigma_Z^2}{N-1}. \quad (4.8)$$

4.3 Importance Sampling

The most simple implementation of Monte Carlo integration is to homogeneously evaluate the integral in Eq. (4.1) by uniformly sampling over the integration volume. This method can lead to large statistical fluctuations if a significant contribution to the integral is over a small region of the integration volume: only a relatively small number of configurations sample the function in this region. Ideally, we would like the majority of configurations to sample the region where $|f(\mathbf{R})|$ is large.

At this time it is convenient to introduce the *importance-function* $g(\mathbf{R})$ (defined on the same domain as $f(\mathbf{R})$), and rewrite the integral in the form

$$I = \int d\mathbf{R} f(\mathbf{R}) = \int d\mathbf{R} g(\mathbf{R}) \left(\frac{f(\mathbf{R})}{g(\mathbf{R})} \right). \quad (4.9)$$

The best choice of $g(\mathbf{R})$ is such that the quantity $h(\mathbf{R}) = f(\mathbf{R})/g(\mathbf{R})$ is roughly constant; consequentially the fluctuations are greatly diminished. If the importance function has the following properties:

1. $g(\mathbf{R}) \geq 0$
2. Normalization $\int d\mathbf{R} g(\mathbf{R}) = 1$

$g(\mathbf{R})$ can be interpreted as a probability distribution function. In this case the integral in Eq. (4.9) is the expectation value of $h(\mathbf{R})$ with respect to the probability distribution $g(\mathbf{R})$ with

$$I = \langle h \rangle = \bar{h} = \int d\mathbf{R} g(\mathbf{R}) h(\mathbf{R}) \quad (4.10)$$

Our goal is to come up with a method to correctly sample $g(\mathbf{R})$ in the relevant domain to generate a set of configurations $\{\mathbf{R}\}$, such a method is described in the next section. Consequentially, Monte Carlo evaluation of the integral is equivalent to taking the mean of the function with respect to the set $\{\mathbf{R}\}$:

$$h_N = \frac{1}{N} \sum_{i=1}^N h(\mathbf{R}_i), \quad (4.11)$$

where all the weights are unity since the configurations have a stationary distribution $g(\mathbf{R})$. Eq. (4.11) is known as an *estimator* and the CLT (Section 4.2) determines the error of h_N .

4.4 Markov Chains and the Metropolis Method

A way to generate a set of configurations $\{\mathbf{R}\}$ with a stationary distribution defined by the probability density function $\mathcal{P}(\mathbf{R})$ is by a Markov chain [7, 106]. Each new configuration in a Markov chain is generated by a probability distribution which depends on the current state of the system, independent of time and history of the state, *i.e.* it does not depend on how it got there. The entire set of configurations forms a chain. Markov chains are defined in terms of a state-to-state transition probability $\mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}')$. We are interested in $\mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}')$ with the normalization property

$$\sum_{\mathbf{R}} \mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}') = 1, \quad (4.12)$$

where the sum runs over all states of the system. For a process to be Markovian it must meet the requirements of *ergodicity*: all states of the system must be accessible from every other state within a finite number of steps and there can be no periodicity [106].

The Metropolis method [141] is a recipe to generate a Markov chain for the invariant distribution $\mathcal{P}(\mathbf{R})$ by a random walk in configuration space. As a motivation for the Metropolis method, define the quantity $\mathcal{P}(\mathbf{R}, t)$, which is the probability of the system to be at the con-

figuration \mathbf{R} at time t . The rate of change of this probability distribution from t to $t + \Delta t$ is determined by two processes:

1. Configurations moving from \mathbf{R} to other \mathbf{R}' , decreasing $\mathcal{P}(\mathbf{R})$
2. Other configurations moving from \mathbf{R}' to \mathbf{R} , increasing $\mathcal{P}(\mathbf{R})$.

The mathematical relation for the rate of change of $\mathcal{P}(\mathbf{R}, t)$ is given by

$$\frac{\mathcal{P}(\mathbf{R}, t + \Delta t) - \mathcal{P}(\mathbf{R}, t)}{\Delta t} = \sum_{\mathbf{R}'} \mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}') \mathcal{P}(\mathbf{R}', t) - \sum_{\mathbf{R}'} \mathcal{T}(\mathbf{R}' \leftarrow \mathbf{R}) \mathcal{P}(\mathbf{R}, t). \quad (4.13)$$

For the equilibrium or steady-state solution we demand that the rate of change of the probability is zero $\mathcal{P}(\mathbf{R}, t + \Delta t) = \mathcal{P}(\mathbf{R}, t)$, or

$$\sum_{\mathbf{R}'} \mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}') \mathcal{P}(\mathbf{R}') = \sum_{\mathbf{R}'} \mathcal{T}(\mathbf{R}' \leftarrow \mathbf{R}) \mathcal{P}(\mathbf{R}), \quad (4.14)$$

where \mathcal{P} is now independent of t . A particular solution to this equation is the detailed balance condition

$$\mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}') \mathcal{P}(\mathbf{R}') = \mathcal{T}(\mathbf{R}' \leftarrow \mathbf{R}) \mathcal{P}(\mathbf{R}), \quad (4.15)$$

which relates the pair of configurations \mathbf{R} and \mathbf{R}' . In other words, the rate of walkers moving from \mathbf{R} to \mathbf{R}' is equal to the rate of walkers moving from \mathbf{R}' to \mathbf{R} [106].

The next step is to split the transition probability into the product of a *trial-step probability* \mathcal{W} and an *acceptance probability* \mathcal{A} :

$$\mathcal{T}(\mathbf{R} \leftarrow \mathbf{R}') = \mathcal{W}(\mathbf{R} \leftarrow \mathbf{R}') \mathcal{A}(\mathbf{R} \leftarrow \mathbf{R}'). \quad (4.16)$$

The trial-step probability has the normalization property $\sum_{\mathbf{R}} \mathcal{W}(\mathbf{R} \leftarrow \mathbf{R}') = 1$ and the acceptance \mathcal{A} is a number between 0 and 1. For the case where the trial-step probability is symmetric $\mathcal{W}(\mathbf{R} \leftarrow \mathbf{R}') = \mathcal{W}(\mathbf{R}' \leftarrow \mathbf{R})$, the detailed balance condition can now be written

as

$$\frac{\mathcal{A}(\mathbf{R} \leftarrow \mathbf{R}')}{\mathcal{A}(\mathbf{R}' \leftarrow \mathbf{R})} = \frac{\mathcal{P}(\mathbf{R})}{\mathcal{P}(\mathbf{R}')} = q, \quad (4.17)$$

with the solution

$$\mathcal{A}(\mathbf{R} \leftarrow \mathbf{R}') = \begin{cases} q, & \text{if } \mathcal{P}(\mathbf{R}) < \mathcal{P}(\mathbf{R}') \\ 1, & \text{if } \mathcal{P}(\mathbf{R}) > \mathcal{P}(\mathbf{R}') \end{cases} \quad (4.18)$$

These two conditions can be concisely written as $\mathcal{A}(\mathbf{R} \leftarrow \mathbf{R}') = \min[1, q]$.

The *generalized* Metropolis Algorithm is the case where $\mathcal{W}(\mathbf{R} \leftarrow \mathbf{R}') \neq \mathcal{W}(\mathbf{R}' \leftarrow \mathbf{R})$; in this case the detailed balance condition can be written as

$$\frac{\mathcal{A}(\mathbf{R} \leftarrow \mathbf{R}')}{\mathcal{A}(\mathbf{R}' \leftarrow \mathbf{R})} = \frac{\mathcal{W}(\mathbf{R}' \leftarrow \mathbf{R})\mathcal{P}(\mathbf{R})}{\mathcal{W}(\mathbf{R} \leftarrow \mathbf{R}')\mathcal{P}(\mathbf{R}')} = q. \quad (4.19)$$

with the solution

$$\mathcal{A}(\mathbf{R} \leftarrow \mathbf{R}') = \min[1, q]. \quad (4.20)$$

For an example of the generalized Metropolis Algorithm see Algorithm 1.

Algorithm 1 The Generalized Metropolis Algorithm. In order to accept a move with probability q , where $q \in [0, 1]$, choose a random number ζ from the uniform distribution $[0, 1]$. If $\zeta < q$, accept the move, otherwise the move is rejected. Note that in the Metropolis method, if a proposed move is rejected, the configuration returns to its previous state. This algorithm can be performed with a single configuration (walker) or an ensemble of walkers.

```

for all Configurations  $\mathbf{R}_i$  advance to  $\mathbf{R}_{i+1}$  do
  Move from  $\mathbf{R}_i$  to  $\mathbf{R}'_i$  with a probability  $\mathcal{W}(\mathbf{X}'_i \leftarrow \mathbf{X}_i)$ 
  Evaluate the acceptance ratio
   $\mathcal{A}(\mathbf{X}'_i \leftarrow \mathbf{X}_i) = \min \left[ 1, \frac{\mathcal{W}(\mathbf{X}_i \leftarrow \mathbf{X}'_i)\mathcal{P}(\mathbf{R}'_i)}{\mathcal{W}(\mathbf{X}'_i \leftarrow \mathbf{X}_i)\mathcal{P}(\mathbf{R}_i)} \right]$ 
  Compare  $\mathcal{A}(\mathbf{X}'_i \leftarrow \mathbf{X}_i)$  to a random variable  $\zeta \in [0, 1]$ .
  if  $\zeta < \mathcal{A}$  then
    Accept the move,  $\mathbf{R}'_i$  is the new configuration:
     $\mathbf{R}_{i+1} = \mathbf{R}'_i$ 
  else
    Reject the move, stay at  $\mathbf{R}_i$  :
     $\mathbf{R}_{i+1} = \mathbf{R}_i$ 
  end if
end for

```

4.5 Statistics in Monte Carlo

The material in this section closely follows Ref. [142]. The definitions of the mean and variance are,

$$\langle A \rangle = \lim_{M \rightarrow \infty} \left[\frac{1}{M} \sum_{i=1}^M A_i \right] \quad (4.21)$$

and

$$\sigma_A^2 = \langle (A - \langle A \rangle)^2 \rangle, \quad (4.22)$$

respectively. It is convenient to evaluate the variance as

$$\sigma_A^2 = \langle A^2 \rangle - \langle A \rangle^2 \quad (4.23)$$

which can be shown to be equivalent to Eq. (4.22).

In practice Monte Carlo simulations produce a finite number of data points $\{A_i : i = 1, \dots, M\}$ (usually printed to an output file), which requires the mean and variance to be evaluated by the approximate relations

$$\langle A \rangle \approx \frac{1}{M} \sum_{i=1}^M A_i \quad (4.24)$$

and

$$\sigma_A^2 \approx \frac{1}{M-1} \sum_{i=1}^M (A_i - \langle A \rangle)^2. \quad (4.25)$$

The error of the mean follows from Eq. (4.8)

$$\sigma = \sqrt{\frac{\sigma_A^2}{M}}. \quad (4.26)$$

It is always necessary to *equilibrate* the data before doing any statistical analysis. Equilibration entails removing the first block of data points - this is necessary because it takes several Monte Carlo steps before equilibrium (detailed balance) is reached.

Eq. (4.26) assumes that successive values of the random quantity A_i are statistically independent. In reality successive values of A_i are printed to file with sufficient frequency such that they are highly *correlated* with each other. At this point it is necessary to define the correlation function

$$c_A(k) = \frac{\sum_{i=1}^{M-k} (A_i - \langle A \rangle)(A_{i+k} - \langle A \rangle)}{\sum_{i=1}^{M-k} (A_i - \langle A \rangle)^2}, \quad (4.27)$$

which measures the correlation between points k steps apart. The function $c_A(k)$ has the property $c_A(0) = 1$ and decays exponentially with respect to k ; there exists a k_0 such that $c_A(k) \approx 0$ for all $k \geq k_0$. The auto correlation length is defined as

$$\kappa_A = 1 + 2 \sum_{k=1}^{k_0} c_A(k). \quad (4.28)$$

The correct expression for the error of the mean is

$$\sigma = \sqrt{\frac{\sigma_A^2 \kappa_A}{M}}, \quad (4.29)$$

this is known as the error bar. It is always necessary to include the error bar for any Monte Carlo result. The error bar can be written in two ways: for a Monte Carlo simulation with a mean of 0.1267 and an error bar of 0.023 the result can be written as 0.1267 ± 0.0023 or more commonly 0.1267(23).

4.5.1 Data blocking

Data blocking [142, 7] is a method where successive values of A_i are grouped into blocks defined by a block length:

$$\langle A \rangle = \frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{i=(j-1)m+1}^{jm} A_i \right) = \frac{1}{n} \sum_{j=1}^n \langle A \rangle_j. \quad (4.30)$$

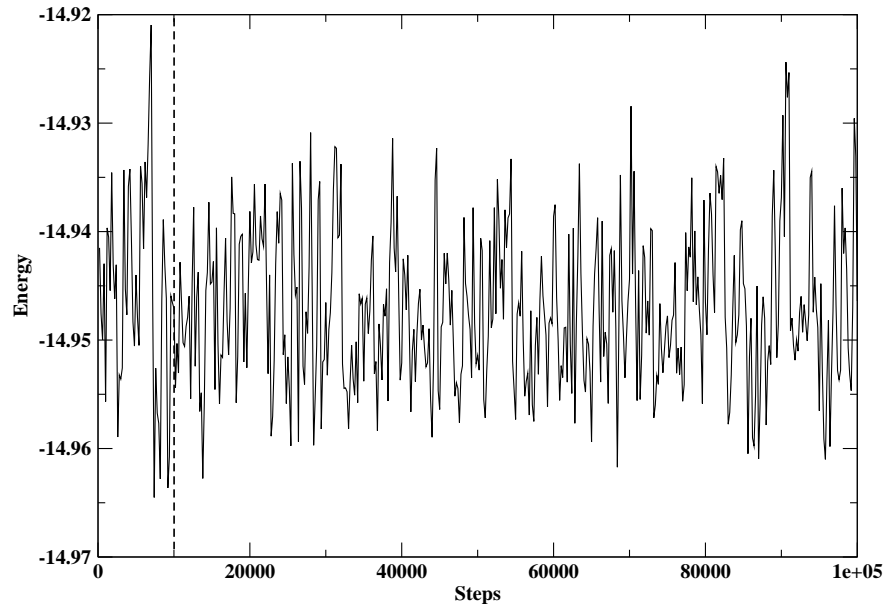
The value $\langle A \rangle_j$ is a block average over all the data between $(j-1)m+1$ and jm , where m is the block size. In the case where the block size m is greater than the autocorrelation length, $m \geq k_0$, the quantities $\langle A \rangle_j$ and $\langle A \rangle_{j+1}$ are statistically uncorrelated. In practice it is not necessary to choose $m \geq k_0$ since this may result in large blocks sizes and consequentially reduce the amount of data available from the Monte Carlo simulation for statistical analysis.

Data blocking can either be performed directly by the Monte Carlo program itself, where only the block averages are printed to the output file, or by a data analysis program. Much of the statistical analysis performed for the work in this thesis is from DataSpork [143]. This is a java based program with a graphical interface that allows the user to define the equilibration length and perform data blocking.

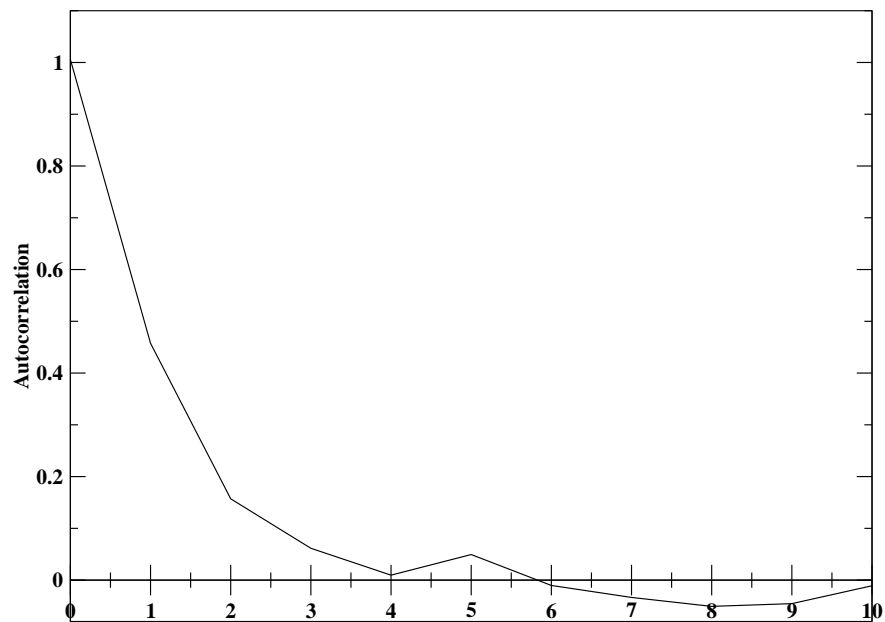
Figure 4.1 shows the results for the local energy of the Li_2 molecule from a variational Monte Carlo (VMC) simulation. The VMC method is discussed in detail in Section 4.7. Figure 4.1(a) plots the local energy with respect to the number of Monte Carlo steps while Figure 4.1(b) plots the autocorrelation function of the data.

4.6 Introduction to Quantum Monte Carlo

Quantum Monte Carlo (QMC) [144, 145, 146, 147, 148, 149, 150, 151, 152, 30, 153, 8, 7, 31, 154] is a many-body method which has been applied to a wide variety of strongly interacting electronic systems. The two QMC techniques discussed in this thesis are variational Monte Carlo (VMC) and the projection based diffusion Monte Carlo (DMC) in the fixed-node approximation. Both of these methods are variational and are only applicable to quantum systems at zero-temperature. VMC and DMC have been able to attain chemical accuracy (about 0.04 eV per molecule) in small systems. In addition, the computational effort of QMC techniques scales with respect to the number of electrons, N , as approximately N^3 , which is advantageous over other many-body methods.



(a) The trace of the local energy.



(b) The autocorrelation of the local energy.

Figure 4.1: 4.1(a) The results of a VMC simulation for the local energy of an Li_2 dimer. The local energy is computed as an integral over $3N$ -dimensions, where $N = 6$ is the number of electrons. The vertical bar denotes the equilibration length; all averages are computed after this cutoff. 4.1(b) The autocorrelation function of the local energy. The autocorrelation length is approximately equal to 4.0; choosing a block size of this length reduces the correlation between successive values of the local energy to zero.

4.7 Variational Monte Carlo

The variational Monte Carlo (VMC) method is based on the combination of the variational theorem from quantum mechanics and the Monte Carlo evaluation of integrals. The integral under consideration is the expectation value of the energy for a trial wave function

$$E[\Psi_T] = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) H \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})}, \quad (4.31)$$

where \mathbf{R} is a $3N$ -dimensional configuration vector representing the coordinates of all the N -electrons and $\Psi_T(\mathbf{R})$ is a many-body trial wave function that explicitly includes correlation. This thesis is strictly limited to real wave functions; therefore $\Psi_T = \Psi_T^*$. This integral can be written in a form more convenient for Monte Carlo

$$E[\Psi_T] = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}) E_L(\mathbf{R})}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})} \quad (4.32)$$

where the local energy $E_L(\mathbf{R})$ is defined as

$$E_L(\mathbf{R}) = \Psi_T^{-1}(\mathbf{R}) H \Psi_T(\mathbf{R}). \quad (4.33)$$

The Metropolis algorithm is used to generate a set of configurations with the probability distribution $\mathcal{P}(\mathbf{R}) = \Psi_T^2(\mathbf{R}) / \int d\mathbf{R} \Psi_T^2(\mathbf{R})$: the energy $E[\Psi_T]$ is estimated by averaging the local energy over this set of configurations:

$$E[\Psi_T] \approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i), \quad (4.34)$$

where equality holds in the limit $M \rightarrow \infty$. The variational energy provides an *upper bound* to the exact ground-state energy

$$E_0 \leq E[\Psi_T], \quad (4.35)$$

where E_0 is the exact ground-state energy.

The specific form of the trial wave function is of the Slater-Jastrow form, and is discussed in detail in Chapter 6. The local energy is constant if the trial wave function Ψ_T is an exact eigenfunction of the Hamiltonian. If $\Psi_T = \Phi_i$, where Φ_i is an eigenfunction of H with the eigenvalue E_i such that $H\Phi_i = E_i\Phi_i$, it can easily be shown that $E_L = E_i$. Consequentially, the variance of the integrand in Eq. (4.32) goes to zero. This is known as the *zero variance* principle for the energy estimator, and applies to the ground-state as well as excited states.

The expectation value of any quantum mechanical operator, \mathcal{A} , diagonal in position space, can be calculated using a similar expression to Eq. (4.32)

$$\langle \mathcal{A} \rangle = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}) \mathcal{A}(\mathbf{R})}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})}, \quad (4.36)$$

where $\mathcal{A}(\mathbf{R}) = \Psi_T(\mathbf{R})^{-1} \mathcal{A} \Psi_T(\mathbf{R})$. The variance of the estimator is given by

$$\sigma_{\mathcal{A}}^2 = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}) [\mathcal{A}(\mathbf{R}) - \langle \mathcal{A} \rangle]^2}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})}. \quad (4.37)$$

In *simple* VMC the trial moves are sampled from a Gaussian centered on the current position of the walker. The size of the move determines the acceptance ratio. The more widely used *importance-sampled* VMC is discussed in Section 4.10.1. For our importance-sampled VMC algorithm see Section 5.2.

Table 4.1 features test results for VMC, where VMC total energies are compared to Hartree-Fock (HF) total energies. The VMC trial wave function is identical to the resulting HF wave function from the SCF calculation. Since the Hartree-Fock approximation contains no correlation, the VMC total energies should agree with the Hartree-Fock total energies to within an error bar, which is confirmed by Table 4.1. We use both an all-electron move and a particle-by-particle move, see Section 5.2 for a more detailed discussion.

Table 4.1: Test results for VMC featuring all-electron move unless otherwise specified. The trial wave function for VMC is identical to the HF wave function and includes no Jastrow correlation factor. All HF wave functions and energies from Ref. [5] unless otherwise specified. All units in a.u.

[†] Particle-by-particle move.

[‡] HF results from Ref. [6].

	HF	VMC	time-step	Accept(%)
Be(¹ <i>S</i>)	-14.57302	-14.57288(31)	0.05	59
N(⁴ <i>S</i>)	-54.40092	-54.40083(34)	0.015	58
N(⁴ <i>S</i>) [†]	-54.40092	-54.40099(24)	0.02	85
O(³ <i>P</i>)	-74.80937	-74.80892(60)	0.0125	53
H ₂ O [‡]	-75.9693	-75.9691(14)	0.01	60

4.8 Umbrella or Correlated Sampling

From Section 4.7, the expectation value of the energy of a trial function Ψ_T can be computed by averaging the local energy $E_L = \Psi_T^{-1} H \Psi_T$ over a set of configurations drawn from the distribution Ψ_T^2 . However, under certain circumstances it may be more convenient to sample the distribution Ψ_G^2 , where Ψ_G is a guiding function, rather than Ψ_T^2 . By re-weighting it is possible to obtain the average with respect to the original distribution Ψ_T^2 :

$$\langle E \rangle = \frac{\int d\mathbf{R} \Psi_G^2(\mathbf{R}) w(\mathbf{R}) E_L(\mathbf{R})}{\int d\mathbf{R} \Psi_G^2(\mathbf{R}) w(\mathbf{R})} = \lim_{M \rightarrow \infty} \frac{\sum_{m=1}^M w(\mathbf{R}_m) E_L(\mathbf{R}_m)}{\sum_{m=1}^M w(\mathbf{R}_m)}, \quad (4.38)$$

where $E_L(\mathbf{R}) = \Psi_T^{-1}(\mathbf{R}) H \Psi_T(\mathbf{R})$ and the weights

$$w(\mathbf{R}) = \frac{\Psi_T^2(\mathbf{R})}{\Psi_G^2(\mathbf{R})} \quad (4.39)$$

convert the distribution Ψ_G^2 to Ψ_T^2 . Some care must be taken in the choice of Ψ_G to insure that the weights are stable. Ideally Ψ_G should either be non-zero over the entire domain, or have a nodal surface close to the nodal surface of Ψ_T . If Ψ_G is computationally less intensive to evaluate than Ψ_T , Eq. (4.38) can represent an improvement in the *efficiency* of VMC.

The efficiency is the error bar achieved in a given amount of computation time. Mathe-

matically this is given by the expression

$$\zeta = \frac{1}{\sigma_\mu^2 T_{CPU}}, \quad (4.40)$$

where σ_μ is the error of the mean and T_{CPU} is the computation time. Typically it is more interesting to compare the efficiency between two calculations

$$\frac{\zeta[1]}{\zeta[2]} = \frac{\sigma_\mu^2[2] T_{CPU}[2]}{\sigma_\mu^2[1] T_{CPU}[1]}. \quad (4.41)$$

So it is possible to improve the relative efficiency by either reducing σ_μ or T_{CPU} or both. The error of the mean is determined by the accuracy of the trial wave function, *i.e.* the inherent variance, and the number of Monte Carlo steps (see Eq. (4.29)). The computation time is determined by the complexity of the trial wave function.

Define the wave functions Ψ_1 and Ψ_2 and their corresponding local energies, $E_L^{(1)} = \Psi_1^{-1} H \Psi_1$ and $E_L^{(2)} = \Psi_2^{-1} H \Psi_2$, respectively. It is possible to calculate the energy difference $\langle \Delta E \rangle = \langle E_L^{(1)} \rangle_{\Psi_G^2} - \langle E_L^{(2)} \rangle_{\Psi_G^2}$ directly by application of Eq. (4.38):

$$\langle \Delta E \rangle = \lim_{M \rightarrow \infty} \frac{\sum_{m=1}^M w^{(1)}(\mathbf{R}_m) E_L^{(1)}(\mathbf{R}_m)}{\sum_{m=1}^M w^{(1)}(\mathbf{R}_m)} - \frac{\sum_{m=1}^M w^{(2)}(\mathbf{R}_m) E_L^{(2)}(\mathbf{R}_m)}{\sum_{m=1}^M w^{(2)}(\mathbf{R}_m)} \quad (4.42)$$

where Ψ_G is the guiding function. Since the energies $E_L^{(1)}(\mathbf{R}_m)$ and $E_L^{(2)}(\mathbf{R}_m)$ are computed over the same set of configurations $(\mathbf{R}_1, \dots, \mathbf{R}_M)$, the energy difference in Eq. (4.42) will have a lower variance than if the two energies were computed separately. This method of computing energy differences is known as correlated sampling [155, 7].

For the guiding function given by the expression [156, 157]

$$\Psi_G^2(\mathbf{R}) = \sum_{i=1}^N \Psi_i^2(\mathbf{R}) \quad (4.43)$$

the weights are

$$w^{(j)}(\mathbf{R}) = \frac{\Psi_j^2(\mathbf{R})}{\Psi_G^2(\mathbf{R})} = \sum_{i=1}^N \left(\frac{\Psi_j^2(\mathbf{R})}{\Psi_i^2(\mathbf{R})} \right) = \left(\sum_{i=1}^N \frac{\Psi_i^2(\mathbf{R})}{\Psi_j^2(\mathbf{R})} \right)^{-1}. \quad (4.44)$$

The set $\{\Psi_i\}$ can represent perturbations from an initial wave function Ψ_0 [7]. Examples of such perturbations can be external electric fields or changes in the geometry of the ions. Another possibility is that the $\{\Psi_i\}$ can represent the ground and excited states of a system. The only restriction on Eq. (4.43) is that all the $\{\Psi_i\}$ are of the same spin-configuration: the dimensions of the determinants for the spin-up and spin-down electrons are all identical. The weights, Eq. (4.44), can be used to evaluate the ratio

$$\frac{\Psi_G^2(\mathbf{R}')}{\Psi_G^2(\mathbf{R})} = \frac{\sum_{i=1}^N \Psi_i^2(\mathbf{R}')}{\sum_{i=1}^N \Psi_i^2(\mathbf{R})} = \frac{\Psi_0^2(\mathbf{R}') \left(1 + \sum_{i=1}^N \frac{\Psi_i^2(\mathbf{R}')}{\Psi_0^2(\mathbf{R}')} \right)}{\Psi_0^2(\mathbf{R}) \left(1 + \sum_{i=1}^N \frac{\Psi_i^2(\mathbf{R})}{\Psi_0^2(\mathbf{R})} \right)}. \quad (4.45)$$

The gradient is determined by the relation

$$\frac{\nabla \Psi_G(\mathbf{R})}{\Psi_G(\mathbf{R})} = \frac{\sum_{i=1}^N \Psi_i(\mathbf{R}) \nabla \Psi_i(\mathbf{R})}{\Psi_G^2(\mathbf{R})}, \quad (4.46)$$

where ∇ is the $3N$ -dimensional gradient operator. Using Ψ_G from Eq. (4.43) in the correlated sampling framework can represent a significant improvement in efficiency if the wave functions are highly correlated. Regardless of the statistical correlation, it is faster to simultaneously calculate several excited states rather than performing independent calculations [51].

4.9 Introduction to Projection Based Methods

While VMC is an excellent method for investigating properties of trial wave functions and is a systematic improvement over single-body methods such as HF and LDA, the ultimate

Table 4.2: The ground-state energy (E^{GS}), excited state energy (E^{EX}) and the energy difference ($\Delta E = |E^{GS} - E^{EX}|$) for GeH_4 and Ge_2H_6 from independent VMC simulations compared to the correlated sampling results (VMC-CS). Results with the CPP are also included. All energies in a.u.

			E^{GS}	E^{EX}	ΔE
GeH_4	No CPP	VMC	-6.17262(27)	-5.78639(34)	0.38623(43)
		VMC-CS	-6.17250(22)	-5.78667(34)	0.38583(36)
	CPP	VMC	-6.23251(27)	-5.85524(33)	0.37727(43)
		VMC-CS	-6.23231(35)	-5.85508(35)	0.37722(35)
	No CPP	VMC	-11.18961(42)	-10.88881(42)	0.30080(59)
		VMC-CS	-11.18915(48)	-10.88852(62)	0.30063(58)
Ge_2H_6	CPP	VMC	-11.30732(45)	-11.01420(49)	0.29312(67)
		VMC-CS	-11.30782(38)	-11.01502(46)	0.29280(44)

accuracy of the algorithm is limited by the quality of the trial wave function.

Projection based techniques are capable of obtaining the true ground-state properties for the N -electron system. This is possible by using the imaginary-time propagator $e^{-\tau(H-E_T)}$, where τ is an interval in imaginary time, H is the N -electron Hamiltonian and E_T is an adjustable trial energy. Applying the imaginary-time propagator to a trial wave function Ψ_T , which has been expanded in terms of the exact eigenfunctions of the Hamiltonian H , leads to

$$e^{-\tau(H-E_T)}\Psi_T = \sum_{i=0} \Phi_i e^{-\tau(E_i-E_T)} \langle \Phi_i | \Psi_T \rangle, \quad (4.47)$$

where E_i is the eigenvalue of the eigenfunction Φ_i . In the large τ limit

$$\lim_{\tau \rightarrow \infty} e^{-\tau(H-E_T)}\Psi_T = \Phi_0 e^{-\tau(E_0-E_T)} \langle \Phi_0 | \Psi_T \rangle, \quad (4.48)$$

the $i = 0$ term dominates if $E_T = E_0$; the terms in the expansion involving the excited states are exponentially damped.¹

The only projection based technique discussed in this thesis is diffusion Monte Carlo

¹It is assumed there is a non-zero overlap between the trial wave function and the ground-state Φ_0 .

(DMC). Examples of other projection based methods are various Green's Function Monte Carlo (GFMC) techniques [146, 158, 151, 159, 7] and Reptation Monte Carlo (RMC) [160, 161].

4.10 Diffusion Monte Carlo

The diffusion Monte Carlo (DMC) method can be derived from the imaginary-time Schrodinger equation

$$-\frac{\partial \Phi(\mathbf{R}, t)}{\partial t} = (H - E_T)\Phi(\mathbf{R}, t), \quad (4.49)$$

where E_T is an energy shift and $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ is a $3N$ -dimensional configuration vector. The solution to Eq. (4.49) is given by

$$|\Phi(t_2)\rangle = e^{-(t_2-t_1)(H-E_T)}|\Phi(t_1)\rangle, \quad (4.50)$$

which can be confirmed by insertion of Eq. (4.50) into Eq. (4.49) and differentiating with respect to t_2 . Let $\tau = t_2 - t_1$ be an interval in imaginary time and $t = t_1$; Eq. (4.50) can be rewritten as

$$|\Phi(t + \tau)\rangle = e^{-\tau(H-E_T)}|\Phi(t)\rangle. \quad (4.51)$$

Eq. (4.51) can be written in integral form by the insertion of the unity operator in the position space representation (the complete set of position states $1 = \int d\mathbf{R}' |\mathbf{R}'\rangle \langle \mathbf{R}'|$) between $\exp[-\tau(H - E_T)]$ and $|\Phi(t)\rangle$, and multiplying on the left by the position state $\langle \mathbf{R}|$:

$$\Phi(\mathbf{R}, t + \tau) = \int d\mathbf{R}' G(\mathbf{R} \leftarrow \mathbf{R}'; \tau) \Phi(\mathbf{R}', t). \quad (4.52)$$

The term $G(\mathbf{R} \leftarrow \mathbf{R}'; \tau)$ is defined as the Green's function for the Hamiltonian H and has the form

$$G(\mathbf{R} \leftarrow \mathbf{R}'; \tau) = \langle \mathbf{R} | e^{-\tau(H-E_T)} | \mathbf{R}' \rangle. \quad (4.53)$$

The Green's function obeys the same imaginary-time Schrodinger equation, Eq. (4.49), as the state $\Phi(\mathbf{R}, t)$:

$$-\partial_\tau G(\mathbf{R} \leftarrow \mathbf{R}'; \tau) = (H(\mathbf{R}) - E_T)G(\mathbf{R} \leftarrow \mathbf{R}'; \tau), \quad (4.54)$$

with the initial condition

$$G(\mathbf{R} \leftarrow \mathbf{R}'; 0) = \delta(\mathbf{R} - \mathbf{R}'). \quad (4.55)$$

The Green's function in Eq. (4.53) cannot be evaluated directly, otherwise there would be no need for DMC, but can be approximated for short time intervals. This is possible by application of the Trotter-Suzuki formula for the exponential of a sum of non-commuting operators \hat{A} and \hat{B}

$$e^{-\Delta\tau(\hat{A}+\hat{B})} = e^{-\Delta\tau\hat{B}/2}e^{-\Delta\tau\hat{A}}e^{-\Delta\tau\hat{B}/2} + \mathcal{O}[\Delta\tau^3], \quad (4.56)$$

where $\Delta\tau$ is a short time interval. The Hamiltonian ($\hat{H} = \hat{T} + \hat{V}$) is the sum of the N -electron kinetic energy operator \hat{T} and a potential \hat{V} , which consists of electron-electron, ion-electron and ion-ion terms. The short time approximation to the Green's function

$$G(\mathbf{R} \leftarrow \mathbf{R}'; \tau) \approx e^{-\Delta\tau[V(\mathbf{R})-E_T]/2} \langle \mathbf{R} | e^{-\Delta\tau\hat{T}} | \mathbf{R}' \rangle e^{-\Delta\tau[V(\mathbf{R}')-E_T]/2}, \quad (4.57)$$

from Appendix A.6 it can be shown that Eq. (4.57) yields

$$G(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) \approx (2\pi\Delta\tau)^{-3N/2} e^{-(\mathbf{R}-\mathbf{R}')^2/2\Delta\tau} e^{-\Delta\tau[V(\mathbf{R})+V(\mathbf{R}')-2E_T]/2}. \quad (4.58)$$

Projection of the ground-state from the initial distribution requires the long time limit $t \rightarrow \infty$, but the Green's function can only be evaluated for small time intervals. The question arises as to how to reconcile these conditions. The answer is that the Green's function can

be factored into many short time intervals and iterated

$$e^{-\tau H} = \prod_{i=1}^N \left[(e^{-\tau H})^{1/N} \right]_i = \prod_{i=1}^N [e^{-\Delta\tau H}]_i, \quad (4.59)$$

where $\Delta\tau = \tau/N$ is a short time interval.

To illustrate the symmetry between the imaginary-time Schrodinger equation and the *diffusion equation*, neglect the potential $V - E_T$ term in the Hamiltonian. The imaginary-time Schrodinger equation, Eq. (4.49), is a $3N$ -dimensional diffusion equation

$$\frac{\partial \Phi(\mathbf{R}, t)}{\partial t} = \lambda \sum_i \nabla_i^2 \Phi(\mathbf{R}, t), \quad (4.60)$$

where $\lambda = 0.5$ is the diffusion constant. The function $\Phi(\mathbf{R}, t)$ can be interpreted as the probability distribution of walkers diffusing in time and space. The normalized Green's function for the diffusion equation

$$G_{Diffusion}(\mathbf{R} \leftarrow \mathbf{R}'; t) = (4\pi\lambda t)^{-3N/2} e^{-(\mathbf{R}-\mathbf{R}')^2/4\lambda t}, \quad (4.61)$$

is identical to the first term in Eq. (4.58). Eq. (4.60) can be solved stochastically: an initial set of walkers from the distribution $\Phi(t=0)$ is propagated in time using the Green's function, Eq. (4.61), as a transition probability.

The remaining part of the Green's function in Eq. 4.58 is known as the branching factor

$$e^{-\Delta\tau[V(\mathbf{R})+V(\mathbf{R}')-2E_T]/2} = G_B(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) = P. \quad (4.62)$$

This can be interpreted as a time-dependant re-normalization or re-weighting for the diffusion process. In pure DMC the weight of each walker is multiplied the by the branching term after each diffusion step. This method is not efficient since the accumulated product of the weights varies greatly from walker to walker. Usually one walker eventually accumulates all

the weight.

An alternative approach to accumulating weights is the birth/death algorithm, where P is treated as a probability for walkers to advance to the next iteration. If $P < 1$, the walker continues to the next iteration at \mathbf{R} with the probability P . If $P \geq 1$, $INT[P]$ walkers continue at \mathbf{R} and a new walker is created at \mathbf{R} with the probability determined by the remainder $P - INT[P]$, where INT denotes the integer part of a real number. These two cases can be implemented by the command

$$N_{copies} = INT[P + \eta], \quad (4.63)$$

where η is a random number such that $0 \leq \eta \leq 1$, and N_{copies} is the number of walkers continuing to the next iteration at position \mathbf{R} . If $N_{copies} = 0$, then the walker dies.

The diffusion Monte Carlo algorithm described above can be inefficient. The branching process described by Eq. (4.63) can lead to large fluctuations where the potential V diverges. This in turn can lead to large statistical fluctuations in expectation values. The *importance sampling* transformation is a way to reduce these fluctuations [151]. Importance sampling requires the introduction of a *trial* or *guiding* function Ψ_T . The form of the trial function is discussed in more detail in Chapter 6; for now it is assumed that Ψ_T is a good approximation to the ground-state Φ_0 . The trial function is used to bias the stochastic process to produce the distribution $\Psi_T\Phi_0$ in the long time limit, as opposed to Φ_0 in ordinary DMC. Section 4.10.4 discusses how estimators are evaluated with respect to the mixed distribution $\Psi_T\Phi_0$. The importance sampling transformation also plays a crucial role in the fixed node approximation, see Section 4.11.2.

4.10.1 The Importance Sampling Transformation

The importance sampling transformation [147] is to multiply Eq. (4.49) by the trial function Ψ_T . Define the importance function $f = \Psi_T \Phi$, and rearrange the terms to arrive at:

$$-\frac{\partial f(\mathbf{R}, t)}{\partial t} = -\frac{1}{2}\nabla^2 f(\mathbf{R}, t) + \nabla \cdot [\mathbf{V}_D(\mathbf{R})f(\mathbf{R}, t)] + [E_L(\mathbf{R}) - E_T] f(\mathbf{R}, t), \quad (4.64)$$

where $\nabla = (\nabla_1, \nabla_2, \dots, \nabla_N)$ is the $3N$ -dimensional gradient and $E_L = \Psi_T^{-1} H \Psi_T$ is the previously defined local energy. The drift-velocity vector \mathbf{V}_D , is related to the gradient of the trial function:

$$\mathbf{V}_D(\mathbf{R}) = \nabla \ln |\Psi_T(\mathbf{R})| = \Psi_T(\mathbf{R})^{-1} \nabla \Psi_T(\mathbf{R}). \quad (4.65)$$

The effect of \mathbf{V}_D is to increase the density of walkers where the magnitude of the trial function $|\Psi_T|$ is large.

The integral equation corresponding to the modified Schrodinger equation in Eq. (4.64) is

$$f(\mathbf{R}, t + \tau) = \int d\mathbf{R}' K(\mathbf{R} \leftarrow \mathbf{R}'; \tau) f(\mathbf{R}', t). \quad (4.66)$$

The function $K(\mathbf{R} \leftarrow \mathbf{R}'; \tau)$ is the modified Green's function and satisfies the relation

$$K(\mathbf{R} \leftarrow \mathbf{R}'; \tau) = \Psi_T(\mathbf{R}) G(\mathbf{R} \leftarrow \mathbf{R}'; \tau) \Psi_T^{-1}(\mathbf{R}'), \quad (4.67)$$

where $G(\mathbf{R} \leftarrow \mathbf{R}'; \tau)$ is the Green's function defined in Eq. (4.53). The short-time approximation for the modified Green's function can be represented as the product of a drift-diffusion and a branching process

$$K(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) \approx G_D(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) G_B(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau), \quad (4.68)$$

see Appendix A.6. The diffusion part is given by

$$G_D(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) = (2\pi\Delta\tau)^{-3N/2} e^{-[\mathbf{R}-\mathbf{R}'-\Delta\tau\mathbf{V}_D(\mathbf{R}')]^2/2\Delta\tau}, \quad (4.69)$$

where the drift is defined as $(\Delta\tau\mathbf{V}_D(\mathbf{R}'))$ and $\mathbf{V}_D(\mathbf{R})$ is constant over the short time interval $\Delta\tau$. The branching term is given by

$$G_B(\mathbf{R} \leftarrow \mathbf{R}'; \tau) = e^{-\Delta\tau[E_L(\mathbf{R})+E_L(\mathbf{R}')-2E_T]/2}. \quad (4.70)$$

An immediate observation is that branching term now contains the local-energy E_L instead of the potential V . This is advantageous because for a good trial wave function the local energy surface $E_L(\mathbf{R})$ is a smooth function, in the limit $\Psi_T = \Phi_0$ the local energy is constant $E_L = E_0$ and has zero variance. This is the zero variance principle discussed in Section 4.7. Hence branching with E_L has the effect of reducing the population fluctuations and improving the efficiency of DMC.

The energy shift E_T is chosen to keep the weights close to unity, $E_T \approx E_0$. It is necessary to periodically adjust the value of E_T to keep the average population of walkers roughly constant. A constant E_T can lead to a population explosion (or implosion) for $E_T > E_0$ (or $E_T < E_0$). In qmcPACK [3], E_T is adjusted each iteration: the value of E_T for the next iteration $(i+1)$ is given by the expression

$$E_T(i+1) = E_{est}(i) - \frac{1}{n_G\Delta\tau} \ln \left(\frac{P(i)}{P_0} \right). \quad (4.71)$$

The term $E_{est}(i) = \langle E_L \rangle(i)$ is the current value of the mixed estimator of the local energy (see Eq. (4.89)), while the second term attempts reset the current population $P(i)$ approximately n_G iterations (generations) later to a target population P_0 . Typical values of n_G are on the order of $1/\Delta\tau$. The feedback of the current number of walkers into E_T introduces a systematic bias, referred to as the population control bias. For a more detailed discussion of this bias

and how to remove its effect on estimators see Ref. [8]. The bias can generally be suppressed by using a large number of walkers. Figure 4.2(a) compares the mixed estimator of the local energy to E_T for the Si_5H_{12} nanocluster; Figure 4.2(b) plots the walker population.

4.10.2 Sampling the Drifted Gaussian

Ignoring the branching term $[E_L(\mathbf{R}) - E_T]$, Eq. (4.64) can be written as

$$\frac{\partial}{\partial t} f(\mathbf{R}, t) = \frac{1}{2} \sum_i \nabla_i \cdot [\nabla_i - \mathbf{F}_i(\mathbf{R})] f(\mathbf{R}, t). \quad (4.72)$$

where $\mathbf{F}_i = 2\nabla_i \Psi_T / \Psi_T$ is related to the i th component of the drift-velocity. This is a Fokker-Planck equation [162] for isotropic diffusion in $3N$ -dimensions, with the a steady state solution of $f = \Psi_T^2$.

Fokker-Planck trajectories are generated by a Langevin equation. The Langevin equation for Eq. (4.72) is

$$\frac{\partial \mathbf{R}(t)}{\partial t} = \lambda \mathbf{F}[\mathbf{R}(t)] + \nu, \quad (4.73)$$

where ν is a Gaussian random variable of mean zero and variance 2λ . A discrete form to Eq. (4.73) can be derived by integrating over short time interval Δt :

$$\mathbf{R}' = \mathbf{R} + \lambda \mathbf{F}[\mathbf{R}(t=0)] \Delta t + \chi, \quad (4.74)$$

where \mathbf{F} is constant over the time interval, and χ is a Gaussian of mean zero and variance $2\lambda\Delta t$. The discrete form of the Langevin equation in Eq. (4.74) prescribes a move for a walker from \mathbf{R} to \mathbf{R}' and is useful for Monte Carlo, but is an approximation to the continuous form in Eq. (4.73). The result of using Eq. (4.74) is the introduction of a bias in the trajectories for all time steps $\Delta t > 0$. The bias increases with the step size and can be corrected by the introduction of a Metropolis accept/reject step.

The Green's function for the Fokker-Planck equation

$$K_{FP}(\mathbf{R} \leftarrow \mathbf{R}'; t) = \langle \mathbf{R} | e^{\lambda t (\nabla^2 - \nabla \cdot \mathbf{F} - \mathbf{F} \cdot \nabla)} | \mathbf{R}' \rangle, \quad (4.75)$$

obeys an identical relation to Eq. (4.72) and has the boundary condition $K(\mathbf{R} \leftarrow \mathbf{R}'; t = 0) = \delta(\mathbf{R} - \mathbf{R}')$. By assuming \mathbf{F} is constant over a short time interval Δt (between \mathbf{R}' and \mathbf{R}) it is possible to obtain the relation

$$K_{FP}(\mathbf{R} \leftarrow \mathbf{R}'; \Delta t) \approx (4\pi\lambda\Delta t)^{-3N/2} e^{-(\mathbf{R} - \mathbf{R}' - \lambda\mathbf{F}(\mathbf{R}'))^2 / 4\lambda\Delta t}. \quad (4.76)$$

This is the drift-diffusion part of the Green's function in Eq. (4.69). The move prescribed by Eq. (4.74) is consistent with Eq. (4.76); a drifting Gaussian with variance $\lambda\Delta t$ and drift $2\lambda\nabla(\ln \Psi_T)$.

Going back to the generalized Metropolis method, the move prescribed by Eqs. (4.74, 4.76) can be treated as a trial-step probability and is accepted with the probability

$$A(\mathbf{R} \leftarrow \mathbf{R}') = \min[1, Q(\mathbf{R} \leftarrow \mathbf{R}')], \quad (4.77)$$

where

$$Q(\mathbf{R} \leftarrow \mathbf{R}') = \frac{K_{FP}(\mathbf{R}' \leftarrow \mathbf{R}; \Delta t) \Psi_T(\mathbf{R})^2}{K_{FP}(\mathbf{R} \leftarrow \mathbf{R}'; \Delta t) \Psi_T^2(\mathbf{R}')}. \quad (4.78)$$

4.10.3 The Metropolis Accept/Reject Step

The *exact* importance-sampled Green's function for the modified Schrodinger equation, Eq. (4.64), satisfies the condition

$$K(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) \Psi_T^2(\mathbf{R}') = K(\mathbf{R}' \leftarrow \mathbf{R}; \Delta\tau) \Psi_T^2(\mathbf{R}), \quad (4.79)$$

which can be used to impose detailed balance.² The Green's function $G_D(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)$ approximately satisfies an identical relation to Eq. (4.79), this can be shown by the substitution $K(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) \approx G_D(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)G_B(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)$ and the symmetry of the branching term $G_B(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau) = G_B(\mathbf{R}' \leftarrow \mathbf{R}; \Delta\tau)$. For the trial move from \mathbf{R}' to \mathbf{R} prescribed by $G_D(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)$, detailed balance is imposed by accepting the move with the probability

$$A(\mathbf{R} \leftarrow \mathbf{R}') = \min[1, Q(\mathbf{R} \leftarrow \mathbf{R}')], \quad (4.80)$$

where

$$Q(\mathbf{R} \leftarrow \mathbf{R}') = \frac{G_D(\mathbf{R}' \leftarrow \mathbf{R}; \Delta\tau)\Psi_T(\mathbf{R})^2}{G_D(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)\Psi_T^2(\mathbf{R}')}. \quad (4.81)$$

This step is only necessary because a short time approximation is used for the Green's function K . As $\Delta\tau \rightarrow 0$, $Q \rightarrow 1$ and the approximate Green's function more nearly satisfies the detailed balance condition, Eq. (4.79). The accept/reject step in Eq. (4.80) ensures that the proper distribution is sampled for any $\Delta\tau$. In practice $\Delta\tau$ is chosen such that the acceptance ratio is greater than 99% : rejecting too many moves has been found to bias the unknown distribution $f(\mathbf{R}, \tau)$ [30, 31]. If branching is ignored than the steady state solution is $f = |\Psi_T|^2$, and DMC reduces to VMC.

The Metropolis accept/reject step, Eq. (4.80), allows any transition probability prescribed by $G_D(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)$, provided that the reverse probability is finite [154]. A finite cutoff for the drift velocity \mathbf{V}_D is introduced to prevent divergences which can occur at singularities in the potential or at the nodal surface of Ψ_T (for a discussion of the nodal surface see Section 4.11.1). The smooth cutoff proposed by Umrigar, Nightingale and Runge (UNR) [8]

$$\bar{\mathbf{V}}_D = \frac{-1 + \sqrt{1 + 2V_D^2\Delta\tau}}{V_D^2\Delta\tau} \mathbf{V}_D \quad (4.82)$$

is implemented in qmcPACK [3]. The modified drift velocity $\bar{\mathbf{V}}_D$ reduces to usual drift veloc-

²Eq. (4.79) can be verified by the definition of $K(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)$ in Eq. (4.67) and the symmetry of $G(\mathbf{R} \leftarrow \mathbf{R}'; \Delta\tau)$ upon the interchange of \mathbf{R} and \mathbf{R}' .

ity \mathbf{V}_D for small $V_D^2 \Delta\tau$, however the magnitude of \bar{V}_D is limited to $\sqrt{2/\Delta\tau}$ for large $V_D^2 \Delta\tau$. The presence of a large drift velocity can lead to occurrence of *persistent configurations*. These are defined as configurations which are stuck in the same position in configuration space and can have a catastrophic effect on estimators such as the local energy. Another method that is used to prevent persistent configurations involves assigning an age to each walker. If a walker becomes old - it has not moved in several iterations - its weight is reduced, increasing the likely-hood that it is killed in the branching procedure.

The actual time the electrons drift and diffuse is less than $\Delta\tau$, due to the Metropolis rejection step in Eq. (4.80). For this reason some DMC algorithms, following Refs. [8, 30], define an effective time step. The effective time step $\Delta\tau_{eff}$ is given by the ratio

$$\Delta\tau_{eff} = \Delta\tau \left[\frac{\langle \Delta R^2 \rangle_{accepted}}{\langle \Delta R^2 \rangle_{total}} \right], \quad (4.83)$$

where ΔR are the displacements resulting from the diffusion part of the Green's function and the angular brackets denote an average over all the moves. The effective time step is used in the branching part of the Green's function G_B .

4.10.4 The Evaluation of Estimators

The computation of ground-state expectation values in DMC is complicated by the fact that in the long time limit DMC yields a set of walkers with the probability distribution $f = \Psi_T \Phi_0$, where Φ_0 is the many-body ground-state. At first sight this might be seen as being problematic because ground-state expectations are computed with respect to the ground-state distribution $|\Phi_0|^2$.

The *mixed estimator* is a method to compute the expectation value of an observable with respect to the mixed distribution f :

$$\langle \mathcal{A} \rangle_{DMC} = \frac{\langle \Phi_0 | \mathcal{A} | \Psi_T \rangle}{\langle \Phi_0 | \Psi_T \rangle}, \quad (4.84)$$

which in integral form can be written as

$$\langle \mathcal{A} \rangle_{DMC} = \frac{\int d\mathbf{R} \Phi_0(\mathbf{R}) \mathcal{A}(\mathbf{R}) \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Phi_0(\mathbf{R}) \Psi_T(\mathbf{R})}, \quad (4.85)$$

it is assumed that the operator \mathcal{A} is diagonal in position space. For any operator that commutes with the Hamiltonian H , it can be shown that the mixed estimator is equal to the true quantum mechanical estimator:

$$\langle \mathcal{A} \rangle_{|\Phi_0|^2} = \frac{\langle \Phi_0 | \mathcal{A} | \Phi_T \rangle}{\langle \Phi_0 | \Phi_T \rangle} = \frac{\langle \Phi_0 | \mathcal{A} | \Phi_0 \rangle}{\langle \Phi_0 | \Phi_0 \rangle}. \quad (4.86)$$

The proof of Eq. (4.86) follows from the fact that Φ_0 can be written as a projected state ($\Phi_0 = \lim_{\tau \rightarrow \infty} e^{-\tau H} \Psi_T$) :

$$\begin{aligned} \langle \mathcal{A} \rangle_{DMC} &= \lim_{\tau \rightarrow \infty} \frac{\int d\mathbf{R} \Psi_T(\mathbf{R}) e^{-\tau H} \mathcal{A} \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T(\mathbf{R}) e^{-\tau H} \Psi_T(\mathbf{R})} \\ &= \lim_{\tau \rightarrow \infty} \frac{\int d\mathbf{R} \Psi_T(\mathbf{R}) e^{-\tau H/2} \mathcal{A} e^{-\tau H/2} \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T(\mathbf{R}) e^{-\tau H/2} e^{-\tau H/2} \Psi_T(\mathbf{R})} \\ &= \frac{\int d\mathbf{R} \Phi_0(\mathbf{R}) \mathcal{A} \Phi_0(\mathbf{R})}{\int d\mathbf{R} \Phi_0(\mathbf{R}) \Phi_0(\mathbf{R})}. \end{aligned} \quad (4.87)$$

The commutation of \mathcal{A} and H was necessary to move \mathcal{A} between the two halves of the projector [154].

For the case $\mathcal{A} = H$ it follows

$$\begin{aligned} \frac{\int d\mathbf{R} \Phi_0(\mathbf{R}) H \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Phi_0(\mathbf{R}) \Psi_T(\mathbf{R})} &= \frac{\int d\mathbf{R} \Phi_0(\mathbf{R}) \Psi_T(\mathbf{R}) [\Psi_T^{-1}(\mathbf{R}) H \Psi_T(\mathbf{R})]}{\int d\mathbf{R} \Phi_0(\mathbf{R}) \Psi_T(\mathbf{R})} \\ &= \frac{\int d\mathbf{R} f(\mathbf{R}) E_L(\mathbf{R})}{\int d\mathbf{R} f(\mathbf{R})} \\ &\approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i). \end{aligned} \quad (4.88)$$

Hence the estimator for the local energy $\langle E_L \rangle$ is a measure of the exact ground-state energy

Table 4.3: The HF, VMC and DMC total energy results for several atoms and molecules compared to experiment. Atomic orbitals from Ref. [5]. All experimental (non-relativistic) results quoted from Ref. [7] and references therein. All energies in a.u.

^a bond length 1.401 Bohr.

	HF	VMC	DMC	Expt. Ref [7]
He(¹ <i>S</i>)	-2.86168	-2.88832(84)	-2.90384(23)	-2.9037
Li(² <i>S</i>)	-7.43273	-7.46046(29)	-7.47663(93)	-7.4781
Be(¹ <i>S</i>)	-14.5730	-14.5733(22)	-14.6655(21)	-14.6673
C(³ <i>P</i>)	-37.6886	-37.6908(24)	-37.8318(25)	-37.8451
(H ₂) ^a	-	-1.157934(55)	-1.174383(89)	-1.17447

E_0 .

Expectation values of operators that do not commute with H can be computed by the *extrapolated estimator*

$$\langle \mathcal{A} \rangle_{EXT} = 2\langle \mathcal{A} \rangle_{DMC} - \langle \mathcal{A} \rangle_{VMC} + \mathcal{O}[(\Psi_T - \Phi_0)^2], \quad (4.89)$$

the VMC estimator corrects the mixed estimator from DMC to extrapolate the ground-state estimate. The accuracy of the extrapolated estimator depends on the quality of the trial wave function, *i.e.* how close Ψ_T is to the true ground-state Φ_0 . Examples of operators that do not commute with H are correlation functions, structure factors and probability densities.

An alternative method to evaluate the estimators of operators that do not commute with H is by the *future* or *forward walking* technique, discussed in Ref. [163]. This method is computationally more intensive than evaluating the extrapolated estimator, and has been criticized for being numerically unstable [164]. The forward walking technique is currently not implemented in qmcPACK [3].

Table 4.4: Comparison of QMC results from qmcPACK [3] to the published results of Umrigar, Nightingale and Runge (UNR) [8]. For the trial wave function of Ne, see Table V in Ref. [8]; for the trial wave function of Li₂, see Table VI in Ref. [8]. The Li₂ dimer is oriented along the z -axis. The electron-electron Jastrow correlation is the Pade form and is *spin-dependent*.

	qmcPACK		UNR	
	VMC	DMC	VMC	DMC
Ne	-128.71457(85)	-128.928(3)	-128.713(2)	-128.922(4)
Li ₂	-14.94726(27)	-14.99063(69)	-14.9472(2)	-14.9890(2)

4.11 Diffusion Monte Carlo for Fermion Systems

4.11.1 The Fermion Sign Problem

The discussion of the DMC method has implicitly assumed that the distribution $f = \Psi_T \Phi_0$ is positive over the full domain, such that it can be interpreted as a probability distribution. However, for a systems of interacting fermions this assumption is no longer valid. The particle exchange antisymmetry condition, Eq. (2.14), implies that there must be positive and negative regions of the wave function and hence the Green's function.³

One way to deal with the *fermion-sign problem* is to assign walkers positive or negative signs, but such methods have proven to be unstable [31]. In the widely used *fixed-node approximation* [148, 149, 152, 30], the unknown nodal surface of Φ_0 is chosen to be identical to the nodal surface prescribed by a trial wave function Ψ_T , such that f is non-negative over the full domain [30]. The nodal surface of Ψ_T , which is a function of $3N$ -dimensions, is a $(3N - 1)$ hyper-surface on which Ψ_T is zero and across which it changes sign. The fixed-node DMC (FN-DMC) algorithm succeeds to produce the lowest many-electron state consistent with the nodal surface of Ψ_T . Before examining the fixed-node approximation in more detail it is useful to discuss the more general *fixed-phase approximation* [165].

³It should be noted that fermion statistics do not represent a complication for VMC since $|\Psi_T|^2$ is always positive.

4.11.2 The Fixed-Phase and Fixed-Node Approximations

The fixed-phase approximation [165] is applicable to systems that do not obey *time-reversal symmetry*. The Hamiltonian for a system of N interacting fermions in an external magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$ (in atomic units)

$$H = \frac{1}{2} \sum_i [-i\nabla_i + \alpha\mathbf{A}(\mathbf{r}_i)]^2 + V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N), \quad (4.90)$$

is an example of a Hamiltonian that is not time-reversal invariant. The eigenfunctions of this Hamiltonian are complex. A generic state $\Phi(\mathbf{R})$ can be written in terms of a modulus $|\Phi(\mathbf{R})|$ and a phase $\varphi(\mathbf{R})$:

$$\Phi(\mathbf{R}) = |\Phi(\mathbf{R})| \exp[i\varphi(\mathbf{R})], \quad (4.91)$$

where $|\Phi(\mathbf{R})|$ and $\varphi(\mathbf{R})$ are real functions. The real and imaginary parts of the Schrodinger equation yield, respectively, a pair of *coupled* differential equations for $|\Phi(\mathbf{R})|$ and $\varphi(\mathbf{R})$:

$$\left[-\frac{1}{2} \sum_i \nabla_i^2 + V_{FP}(\mathbf{R}) + V(\mathbf{R}) \right] |\Phi(\mathbf{R})| = E|\Phi(\mathbf{R})|, \quad (4.92)$$

and

$$\sum_i \nabla_i \cdot \{ |\Phi(\mathbf{R})|^2 [\nabla_i \varphi(\mathbf{R}) + \alpha\mathbf{A}(\mathbf{r}_i)] \} = 0, \quad (4.93)$$

where the effective potential V_{FP} is given by

$$V_{FP}(\mathbf{R}) = \frac{1}{2} \sum_i [\nabla_i \varphi(\mathbf{R}) + \alpha\mathbf{A}]^2. \quad (4.94)$$

Eq. (4.92) for $|\Phi(\mathbf{R})|$ is a boson problem, since $|\Phi(\mathbf{R})| \geq 0$, but it is coupled to Eq. (4.93). The fixed-phase approximation consists of choosing an appropriate trial function $\Psi_T(\mathbf{R})$, with the symmetry of the desired many-body ground-state $\Phi_0(\mathbf{R})$, and fixing the unknown phase of $\Phi_0(\mathbf{R})$ to be identical to the phase of the trial function, $\varphi_T(\mathbf{R}) = \Im(\log \Psi_T(\mathbf{R}))$. For the phase determined by $\varphi_T(\mathbf{R})$, the boson problem for $|\Phi(\mathbf{R})|$ is solved exactly by DMC.

The fixed-node approximation is a special case of the fixed-phase approximation, and is applicable when time-reversal symmetry is obeyed, *i.e.* $\mathbf{A} = 0$. The phase of the state can either be 0 or π , corresponding to regions where the wave function is either positive or negative, respectively. The phase changes from the values 0 to π discontinuously at the nodes of the trial function: $\varphi(\mathbf{R}) = \pi[1 - \Theta(\Phi_T(\mathbf{R})/|\Phi(\mathbf{R})|)]$, where Θ is the unit step (Heaviside) function. The corresponding fixed-node potential

$$V_{FP}(\mathbf{R}) = V_{FN}(\mathbf{R}) = \pi^2 (\nabla_i \Psi_T(\mathbf{R}))^2 \delta^2(\Psi_T(\mathbf{R})), \quad (4.95)$$

has the effect of not allowing paths in configuration space that cross the nodal surface. Multiplying Eq. (4.92) by the trial function Ψ_T when $\mathbf{A} = 0$ yields an identical relation to Eq. (4.64): the modified Schrodinger equation for the state $f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R})\Phi_0(\mathbf{R}, \tau)$. f can now be interpreted as a probability distribution since it is positive.

The implementation of the fixed node method within the importance sampling DMC algorithm (see Section 5.3) is straightforward. Any proposed move which crosses the nodal surface is rejected: $K(\mathbf{R} \leftarrow \mathbf{R}'; \tau) = 0$ if \mathbf{R} and \mathbf{R}' are in different nodal pockets. Eq. (4.92) is solved for each nodally bound volume separately; each solution can be viewed as the Bose ground-state for that volume with hard wall boundary conditions on the enclosing nodal surface. The only communication between the different volumes is by the trial energy E_T , see Eq. (4.71), which is adjusted to maintain an average walker population. It should be noted that some algorithms kill walkers that cross a node, but this practice has been found to increase the time step error [8, 31, 166]. qmcPACK by default rejects moves which cross the nodal surface [3].

The result of the stochastic process defined above is the lowest energy many-electron state consistent with the nodal surface of Ψ_T . The accuracy of the method is determined by how closely the nodal surface of Ψ_T approximates the true ground-state nodal surface. The fixed-node DMC energy is always greater than or equal to the true ground-state energy

($E_{FN} \geq E_0$), with equality holding in the limiting case when the nodes of Ψ_T are exact. In this sense fixed-node DMC is a variational method.⁴

The number of walkers attempting to cross the nodes of Ψ_T vanish as $\Delta\tau \rightarrow 0$. For any finite time step $\Delta\tau$, walkers will rarely attempt to cross the nodes due to the approximate nature of the Green's function K . This can be understood by inspection of drift-diffusion Green's function, Eq. (4.69). The drift term, which acts to push walkers away from the nodes, is proportional to $\Delta\tau$, while the standard deviation of the Gaussian from diffusion is proportional to $\sqrt{\Delta\tau}$. Since $\Delta\tau$ is always a small number, $\Delta\tau < 1$ a.u., the diffusion term can occasionally overcome the drift, resulting in a node crossing. The failure of the approximate form of the Green's function is especially pronounced near the nodes where the local-energy and the drift-velocity diverge, and at singularities in the potential (electron-electron and electron-ion overlaps) which can result in divergences of the local energy and a discontinuity of the drift velocity.

4.12 Excited States in QMC

So far almost all of the discussion pertaining to the QMC method has been confined to ground-state properties. The trial wave function has been assumed to be an approximation to the N -electron ground-state: $\Psi_T \Rightarrow |N;0\rangle$. In general we are interested in electronic excitations, which in QMC are calculated as total energy differences. Electronic excitations such as electron addition and removals can be calculated as the total energy difference of the ground-states of the N and $N \pm 1$ systems, *i.e.* $|E_N - E_{N \pm 1}|$. This is the idea behind Δ SCF methods discussed in Section 2.3.2. For particle conserving excitations such as the optical gap, it is necessary to construct the trial wave function for the excited state: $\Psi_T \Rightarrow |N;i\rangle$, where the index i denotes the excited state of the N electron system [167, 168, 31, 45]. Usually such excited state wave functions are created by promoting an electron to an excited

⁴For a proof see Ref. [30].

state orbital in the Slater determinant part. Such approaches have been used for molecules and clusters in Refs. [2, 57, 67, 169, 66]. This approach has also been used in solids to calculate band gaps. Previous QMC calculations of band gaps include atomic solid nitrogen in the $I2_13$ structure [170], diamond [171, 168] and silicon [172].

The VMC method can be implemented with any trial wave function. The zero-variance principle, see Section 4.7, applies to all the eigenfunctions of a given Hamiltonian H , including the ground and excited states. For a trial wave function of a distinct symmetry, the variational principle guarantees that the energy is greater than or equal to the lowest eigenvalue of that symmetry.

At first it might seem that DMC cannot be extended to excited states, since the DMC algorithm projects the lowest energy state of H in the long time limit. As we have already discussed in Section 4.11.2, DMC in the fixed-node approximation converges to the lowest energy many-body state consistent with the nodes of a trial wave function Ψ_T . The implication is that if the nodal surface of Ψ_T is identical to the nodal surface of an eigenfunction of H , DMC will project the exact eigenfunction. Therefore, the fixed-node approximation prevents a collapse to the ground-state. If Ψ_T does not have correct nodal surface of the desired excited state, the DMC energy may not be above exact energy of that state due to mixing of lower energy states.

It was widely believed that the DMC energy must be greater than or equal to the energy of the lowest eigenfunction with the same symmetry as the trial wave function. Foulkes *et al.* [167] showed that this assumption is only justified if the trial wave function transforms according to a one-dimensional irreducible representation of the symmetry group of the Hamiltonian H . If Ψ_T transforms to a multi-dimensional irreducible representation, corresponding to a degenerate eigenfunction, the DMC energy may lie below the energy of the lowest eigenfunction of that symmetry. Despite this, DMC in the FN approximation works well, regardless of the dimension of the irreducible representation.

4.13 Calculating the Optical Gap in QMC

QMC is an excellent approach to calculate optical properties since optical excitations are inherently many-body in character: including the interaction of the electron-hole pair with all of the other electrons. Optical excitations in QMC are determined as the total energy difference between the ground-state configuration and an excited state configuration

$$E^{OPT} = E^{EX} - E^{GS}. \quad (4.96)$$

The disadvantage of this approach is that it becomes difficult for QMC to attain the absorption spectrum of the nanoclusters over a large energy range. Our interest is confined to the optical gap, which is the lowest dipole allowed singlet transition in the absorption spectrum. All of the QMC calculations are for vertical excitations in the ground-state structure of the constituent atoms (see Figure 1.1); structural relaxation of the excited state was not undertaken.

The excited state configuration in Eq. (4.96) is created by removing an electron from the highest occupied molecular orbital (HOMO) and promoting it to an excited state molecular orbital in the spin-up D^\uparrow Slater determinant:

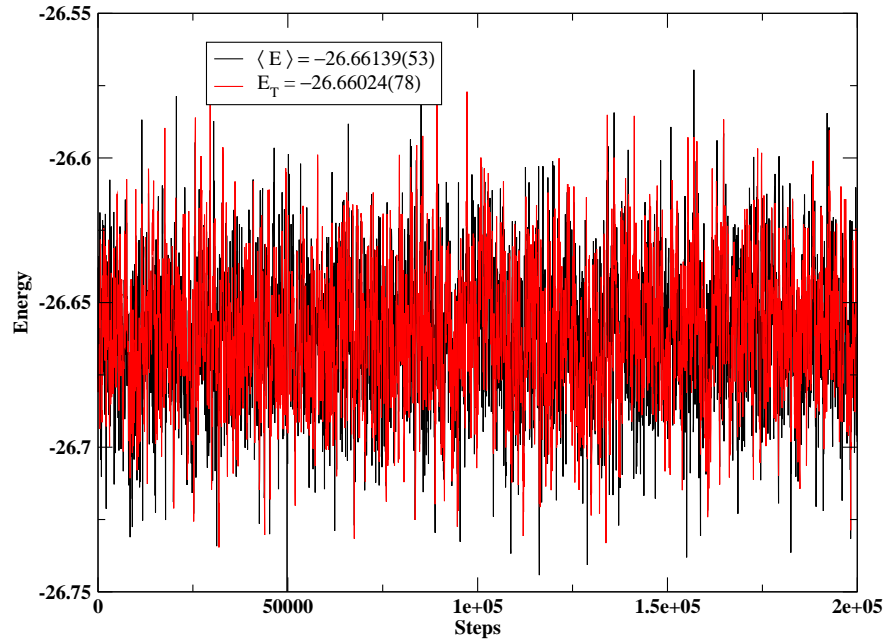
$$D^\uparrow = \begin{vmatrix} \psi_1(\mathbf{r}_1) & \dots & \psi_1(\mathbf{r}_{N/2}) \\ & & \dots \\ \psi_{N/2}(\mathbf{r}_1) & \dots & \psi_{N/2}(\mathbf{r}_{N/2}) \end{vmatrix} \rightarrow \begin{vmatrix} \psi_1(\mathbf{r}_1) & \dots & \psi_1(\mathbf{r}_{N/2}) \\ & & \dots \\ \psi_{N/2+1}(\mathbf{r}_1) & \dots & \psi_{N/2+1}(\mathbf{r}_{N/2}) \end{vmatrix}. \quad (4.97)$$

For the optical transition to be allowed, a dipole matrix element between the HOMO and excited molecular orbital must exist. The excited molecular orbital is normally chosen to be the lowest unoccupied molecular orbital (LUMO), see Eq. (4.97), unless the HOMO \rightarrow LUMO transition is optically forbidden. This was found to be the case for some of the Ge_nH_m and Si_nH_m nanoclusters, such as Ge_5H_{12} .

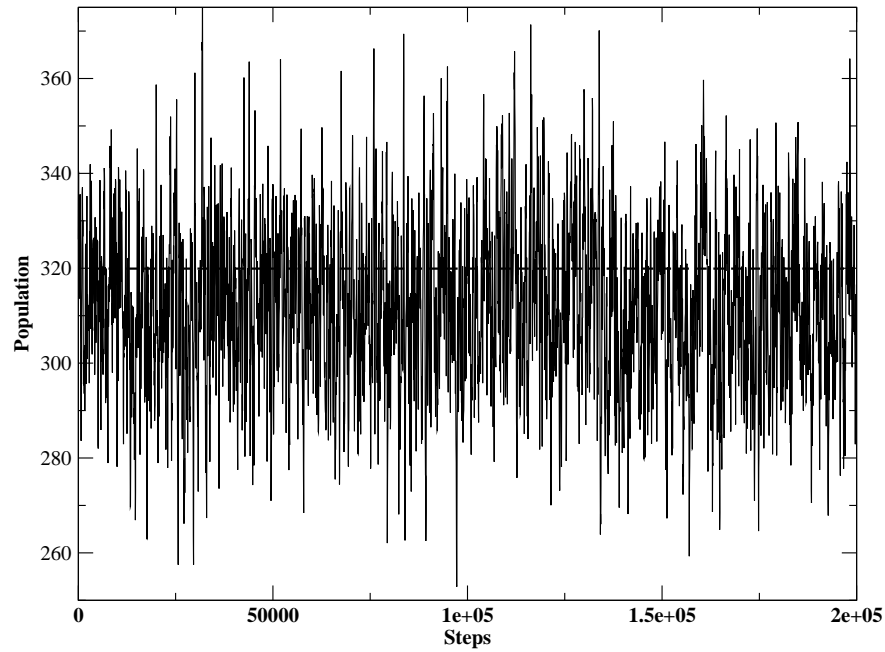
In order to construct the true singlet excited state, it is necessary to add a second set of determinants where the down-spin electron is excited:

$$\frac{1}{\sqrt{2}} \left[\tilde{D}_1^\uparrow D_1^\downarrow + D_2^\uparrow \tilde{D}_2^\downarrow \right], \quad (4.98)$$

where \tilde{D} denotes a Slater determinant where an electron has been excited. The inclusion of the second set of determinants has been found to be a negligible effect to the optical gap [57, 2] and is neglected in this work.



(a) Comparison of the mixed estimator for the local energy $\langle E_L \rangle$ to energy shift E_T .



(b) The FN-DMC population.

Figure 4.2: The results of a Fixed-Node DMC calculation for the Si_5H_{12} cluster. The time-step is 0.025 a.u., the target population is 320 walkers and $n_G = 100$ (see the accompanying text).

Chapter 5

QMC Algorithms and QMCPack

5.1 Introduction

This chapter presents the algorithms for VMC and FN-DMC used in the qmcPACK quantum Monte Carlo package [3]. The development of qmcPACK was one of the most important aspects of the work in this thesis. qmcPACK was written in C++ [173] using the Standard Template Library (STL) and libraries by Boost [138], HDF5 [174], BLAS and LAPACK. For the parallelization we used the MPI package [175]. Previous versions of qmcPACK have used the GNU Scientific Library [176], BLITZ++ [177] and SPRNG [139]. All of the QMC calculations performed for this work are by qmcPACK unless stated otherwise. In addition, we have released the code to the public and are currently working on adding new components, improving algorithms and a more complete manual and documentation.

The VMC and FN-DMC algorithms presented in this chapter are for a particle-by-particle move, in which each of the particles (electrons) are moved individually and either Metropolis accepted/rejected. It is also possible to use an all-electron move, in which all of the electrons of the system are moved simultaneously and the move is Metropolis accepted/rejected. This approach is more closely related to the derivations in Chapter 4. qmcPACK features both types of moves. The advantage of the particle-by-particle move is that it permits a much larger step size to attain a given acceptance ratio. More importantly, for systems containing large number of electrons, it is unfeasible to use an all-electron move.

Caution should always be exercised when performing a QMC simulation featuring a particle-by-particle move on a system with the bare $(-Z/r)$ Coulomb potential. By choosing

a small enough step size, it is possible to attain a reasonable acceptance ratio, but it is often the case that the acceptance rate for moves involving the inner core electrons are much smaller than the acceptance rate for moves involving the outer valence electrons. This is due to the differences in the energy and length scales between the core and valence electrons. Such a disparity between the acceptance rates make any QMC result suspect.

5.2 Variational Monte Carlo Algorithm with Importance Sampling for Particle-by-Particle Move

The algorithm presented in this section is for the importance sampled version of VMC. This algorithm is very similar to the Fixed-Node DMC algorithm presented in the next section; the main difference is that branching is suppressed in VMC. The time-step does not correspond to an interval in imaginary time as in DMC, but is used to control the size of the trial move and therefore the acceptance ratio.

The algorithm incorporates the concept of blocking discussed in Section 4.5.1; the estimators are printed at the end of each block. It should be restated that it is always necessary to equilibrate the data before performing any subsequent data analysis. All the weights in VMC are unity.

for all blocks do

for all steps do

for $iw \in$ walkers do

Reset the variables: $nAcceptTemp$, $nRejectTemp$, $= 0$

for $ip \in$ particles do

For the particle at position $\mathbf{r}_{ip}^{(iw)}$ drift and diffuse: $\mathbf{r}_{ip}'^{(iw)} = \mathbf{r}_{ip}^{(iw)} + \chi + \tau \mathbf{V}_D(\mathbf{r}_{ip}^{(iw)})$,

where χ is a gaussian of mean zero and variance τ .

Accept the move with probability: $A(\mathbf{R}^{(iw)} \rightarrow \mathbf{R}'^{(iw)}; \tau) = \min[1, Q(\mathbf{R}^{(iw)} \rightarrow \mathbf{R}'^{(iw)})]$, where Q is defined in Eq. (4.81).

if The move is accepted **then**

Update the walker properties

$nAcceptTemp++$

else

$nRejectTemp++$

end if

end for

Update the walker properties

end for

$nAccept+ = nAcceptTemp, nReject+ = nRejectTemp$

for all walkers **do**

Accumulate estimator for observable (\mathcal{A}):

$\mathcal{A}_{sum}+ = W^{(iw)}\mathcal{A}^{(iw)}, W_{sum}+ = W^{(iw)}$

end for

end for

Write estimator to file and reset variables: $\mathcal{A}_{sum}, W_{sum} = 0$

end for

5.3 Fixed-Node Diffusion Monte Carlo Algorithm for Particle-by-Particle Move

This section presents the version of the FN-DMC algorithm used in qmcPACK [3]. The algorithm follows the derivations in Chapter 4.

For DMC it is always necessary to start the calculation with an initial set of walkers (ensemble). These walkers come from a previous VMC calculation and sample the distribution

Ψ_T^2 , where Ψ_T is the trial wave function. The FN-DMC algorithm starts from this distribution and iteratively projects the mixed distribution $\Phi_0\Psi_T$, where Φ_0 is the lowest energy state consistent with the nodal surface of Ψ_T . Estimators are calculated from this mixed distribution as in Section 4.10.4.

For each walker it is necessary to store a set of properties, in addition to the coordinates of the N -electrons: \mathbf{R} . The following walker properties are typically stored:

- The local energy: $E_L(\mathbf{R})$
- The square of the local energy: $E_L(\mathbf{R})^2$
- The drift velocity: $V_D(\mathbf{R})$
- The weight: W
- The multiplicity: M
- The Age: Age

The multiplicity is used for branching. In qmcPACK [3] we also store the constituent parts of the local energy.

As already mentioned the DMC algorithm is designed for an ensemble of walkers. Since the walker population is changing throughout the run we represent the walker ensemble by a linked list: for a discussion of linked list see Ref. [178]. A ceiling and a floor are set for the walker population. The ceiling is twice the ideal population while the floor is half the ideal population. If the walker population exceeds the ceiling our algorithm kills the extra walkers. On the other hand, if the walker population is below the floor extra walkers are created. If either of these situations arise it is recommended to re-run DMC with a smaller time-step.

```
for all blocks do
  for all steps do
```

for $iw \in \text{walkers}$ **do**

Reset the variables: $nAcceptTemp$, $nRejectTemp$, $\Delta R_{accepted}^2$, $\Delta R_{total}^2 = 0$

for $ip \in \text{particles}$ **do**

For the particle at position $\mathbf{r}_{ip}^{(iw)}$ drift and diffuse: $\mathbf{r}_{ip}'^{(iw)} = \mathbf{r}_{ip}^{(iw)} + \chi + \tau \mathbf{V}_D(\mathbf{r}_{ip}^{(iw)})$,

where χ is a gaussian of mean zero and variance τ .

$$\Delta R^2 = |\chi + \tau \mathbf{V}_D(\mathbf{r}_{ip}^{(iw)})|^2$$

$$\Delta R_{total}^2 += \Delta R^2$$

if There is a node crossing: $[\Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R}) < 0]$ **then**

Reject the move: $nRejectTemp ++$

else

Accept the move with probability: $A(\mathbf{R}^{(iw)} \rightarrow \mathbf{R}'^{(iw)}; \tau) = \min[1, Q(\mathbf{R}^{(iw)} \rightarrow \mathbf{R}'^{(iw)})]$, where Q is defined in Eq. (4.81).

if The move is accepted **then**

Update the walker properties

$nAcceptTemp ++$

$$\Delta R_{accepted}^2 += \Delta R^2$$

else

$nRejectTemp ++$

end if

end if

end for

if All particle moves rejected **then**

Increment the Age: $Age^{(iw)} ++$

else

Update the walker properties

end if

Calculate the effective time-step: $\tau_{eff} = \frac{\tau \Delta R_{accepted}^2}{\Delta R_{total}^2}$

Calculate the branching probability: $P^{(iw)} = \exp \left[-\tau_{eff}(E_L(\mathbf{R}^{(iw)}) + E_L(\mathbf{R}'^{(iw)}) - 2E_T)/2 \right]$

Reweight the walker: $W^{(iw)}_* = P^{(iw)}$

if The age of the walker is greater than one: $Age^{(iw)} > 1$ **then**

Rescale the branching probability: $P^{(iw)} = \min[1, P^{(iw)}]$

end if

Set the multiplicity: $M^{(iw)} = P^{(iw)} + \xi$, where ξ is a random number $\in [0, 1]$

Determine number walkers that advance to the next generation from the multiplicity: $N_{copies}^{(iw)} = INT[M^{(iw)}]$

if $N_{copies}^{(iw)} = 0$ **then**

Kill the walker

end if

end for

$nAccept+ = nAcceptTemp$, $nReject+ = nRejectTemp$

for all walkers **do**

Accumulate estimator for observable (\mathcal{A}):

$\mathcal{A}_{sum}+ = W^{(iw)}\mathcal{A}^{(iw)}$, $W_{sum}+ = W^{(iw)}$

end for

Perform branching.

Reset the walker weights and multiplicities: $W^{(iw)}$, $M^{(iw)} = 1.0$

return The new population of walkers P

Update the trial energy: E_T , see Eq. (4.71).

end for

Write estimator to file and reset variables: \mathcal{A}_{sum} , $W_{sum} = 0$

end for

5.4 Parallelization of QMC

The desirability of the QMC method is that it is fairly straightforward to parallelize and requires minimal inter-processor communication. Parallelization reduces the load on each processor and reduces the total computation time. In the context of QMC, parallelization means distributing the walkers among several processors. This is important because typical QMC simulations (especially DMC) may require several hundred walkers, with parallelization each processor is only responsible for a fraction of the total number. Parallelization is essential for the systems containing large numbers of electrons (such as the clusters simulated in this thesis); placing a large number of walkers on a single processor is unfeasible. With a parallel version of QMC it is possible to assign a manageable number of walkers on each processor while maintaining a large total walker population.

The most trivial parallelization is to perform independent simulations on each processor and collate the results for statistical analysis. No inter-processor communication is required for this method, although it is necessary to assign different seeds to the random number generator for each of the processors.

qmcPACK [3] also features versions of VMC and DMC that incorporate inter-processor communication. For VMC a fixed number of walkers are assigned to each processor and undergo independent random walks; inter-processor communication is required to collate the results (for estimators such as the local energy) at the end of each block. This method cannot be extended to DMC, which has a fluctuating walker population. The DMC algorithm, courtesy of the branching, depends on the total walker population over all of the individual processors. In addition to the total walker population, the trial energy E_T , which is in part determined by the population, is also a global property. It is necessary to keep track of the global walker population and regulate the number of walkers on each processor. qmcPACK regulates the walkers per node by *swapping* walkers between individual processors: a maximum difference of only one or two walkers is allowed to exist between the nodes. Appropriate

actions are taken if the global walker population becomes too large or too small.

Chapter 6

The Quantum Monte Carlo Trial Wave Function

6.1 Introduction

The accuracy and statistical efficiency of any QMC simulation depends upon the quality of the trial wave function Ψ_T .¹ The requirements on Ψ_T are that it must be continuous and differentiable wherever the potential is finite, and the integral $\int d\mathbf{R} \Psi_T H \Psi_T$ must exist [31]. In order for Ψ_T to describe a system of interacting fermions, it must obey particle exchange anti-symmetry

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N) = -\Psi(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N), \quad (6.1)$$

where $\mathbf{x}_i = \{\mathbf{r}_i, \sigma_i\}$ represents the space and spin coordinates of the i th electron.

A specific form that explicitly includes correlation in a many-body sense is known as the Slater-Jastrow type:

$$\Psi_T(\mathbf{X}) = D(\mathbf{X})e^{-U(\mathbf{X})}, \quad (6.2)$$

where $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ denotes the spin and space variables for the N -electrons, D is a Slater determinant and U is a Jastrow correlation function. The Slater determinant automatically satisfies the particle exchange anti-symmetry requirement when the Slater matrix elements are single particle orbitals, usually obtained from Hartree-Fock (HF) or

¹It is necessary to distinguish between the definitions of *accuracy* and *precision*, and how they pertain to QMC. The accuracy is determined by the quality of the trial wave function - in other words the approximation used to construct the trial wave function. The precision is related to the statistical uncertainty, and is determined by the number of Monte Carlo steps and the statistical correlation between each Monte Carlo step.

Kohn-Sham density functional theory (DFT). The Jastrow [179] factor typically contains electron-ion, electron-electron, and even electron-electron-ion or higher order correlations. The Jastrow factor is positive and multiplicative, hence the nodes of Ψ_T are determined solely by the Slater determinant terms.

For any spin-independent Hamiltonian H , it can be shown that the position space function

$$\Psi_T(\mathbf{R}) = D^\dagger(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) D^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N) e^{-U(\mathbf{R})}, \quad (6.3)$$

gives the same expectation value as the spin-space function $\Psi_T(\mathbf{X})$ in Eq. (6.2) [31]. This is true for any spin-independent operator and eliminates the need to sum over the spin variables. The spin-up electrons, by convention labeled $(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow})$, are treated as being distinguishable from the spin-down electrons $(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N)$. One consequence of using the position space function in Eq. 6.3 is that opposite spin electrons are not antisymmetric upon exchange of their spatial coordinates.

The most general form of the trial wave function in Eq. (6.3) is a linear combination of Slater determinants $\sum_i c_i D_i^\dagger D_i^\downarrow$. Multi-determinant trial wave functions typically arise in Configuration Interaction (CI) calculations, see Section 3.2. It also may be necessary to use a multi-determinant if the state must have a correct total spin $\mathbf{S} = \sum_i \mathbf{S}(i)$ or angular momentum $\mathbf{L} = \sum_i \mathbf{L}(i)$. A single Slater determinant is an eigenfunction of the z -component of the spin, $S_z = \sum_i S_z(i)$, and angular momentum, $L_z = \sum_i L_z(i)$, but not necessarily the total spin or angular momentum.

It is generally desirable to use the simplest possible trial wave function since the most computationally intensive part of any QMC simulation is evaluating the trial wave function and its derivatives. The following sections detail how to correctly evaluate the Slater-Jastrow trial function.

6.2 Evaluating the Slater Determinant

The Slater determinant is the determinant of the Slater matrix and is denoted by

$$D(\mathbf{R}) = \det(\underline{D}(\mathbf{R})), \quad (6.4)$$

where the matrix is underlined to prevent confusion with the determinant. The elements of the Slater matrix for the electrons are from a set of N single-particle orbitals

$$D_{ij}(\mathbf{R}) = \phi_j(r_i). \quad (6.5)$$

We assume that the single particle orbitals are a function of a single coordinate. The Slater matrix can be represented by

$$\underline{D}(\mathbf{R}) = \begin{pmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \cdots & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \cdots & \phi_N(\mathbf{r}_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \cdots & \phi_N(\mathbf{r}_N) \end{pmatrix}, \quad (6.6)$$

for the configuration $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$.

For the purposes of evaluating the Slater determinant [180], it is necessary to define the following quantities:

- A. The minor M_{ij} for the element D_{ij} (of the $N \times N$ matrix) is the determinant of an $(N-1) \times (N-1)$ matrix obtained by removing all the elements of row i and column j .
- B. The co-factor matrix \underline{C} is constructed on an element by element basis from the minor using the simple relation:

$$C_{ij} = (-1)^{i+j} M_{ij}$$

C. The inverse of the matrix \underline{D} obeys the following relation

$$\underline{I} = \underline{D}^{-1} \underline{D}, \quad (6.7)$$

where \underline{D}^{-1} is the inverse matrix. The inverse is related to the *transpose* of the cofactor matrix by the relation

$$(D^{-1})_{ij} = \frac{(C_{ij})^T}{D} = \frac{C_{ji}}{D}, \quad (6.8)$$

where D is the determinant.

Using these definitions, the determinant is evaluated as the sum of the products of the elements of any *row* or *column* of the matrix and the corresponding cofactor

$$D = \sum_{j=1}^N D_{ij} C_{ij} \longrightarrow \text{along row } i \quad (6.9)$$

$$= \sum_{i=1}^N D_{ij} C_{ij} \longrightarrow \text{along column } j. \quad (6.10)$$

In quantum Monte Carlo it is necessary to evaluate the ratio

$$\frac{\Psi(\mathbf{R}')}{\Psi(\mathbf{R})} = \frac{D(\mathbf{R}')}{D(\mathbf{R})} e^{-U(\mathbf{R}') + U(\mathbf{R})}, \quad (6.11)$$

where $\mathbf{R}' = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}'_i, \dots, \mathbf{r}_N)$ differs from $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)$ only in the coordinates of the i th particle: D can either be the up or down determinant depending on the spin of particle i . It is important to note that the underlying Slater matrix $\underline{D}(\mathbf{R}')$ is identical to $\underline{D}(\mathbf{R})$ except for the i th row. Hence the elements of the cofactor matrix \underline{C} along row i remain the same, since C_{ij} is constructed from all the other rows of the Slater matrix except row i ; therefore C_{ij} is not a function of the coordinate \mathbf{r}_i . Using Eqs. (6.8) and (6.10), the

ratio of the determinate at \mathbf{R}' to the determinate at \mathbf{R} is [150]

$$\begin{aligned} q = \frac{D(\mathbf{R}')}{D(\mathbf{R})} &= \frac{\sum_{j=1}^N D_{ij}(\mathbf{R}') C_{ij}}{D(\mathbf{R})} \\ &= \sum_{j=1}^N D_{ij}(\mathbf{R}') D_{ji}^{-1}(\mathbf{R}). \end{aligned} \quad (6.12)$$

The reason for using this relation is that it eliminates the need to repeatedly evaluate the determinant for every Monte Carlo step, which takes an order of $\mathcal{O}[N^3]$ operations. If the move from \mathbf{R} to \mathbf{R}' is accepted (particle i), the inverse Slater matrix \underline{D}^{-1} must be updated, which only requires $\mathcal{O}[N^2]$ operations. The matrix \underline{D}^{-1} is actually the inverse of the transpose of \underline{D} . The new inverse Slater matrix (at \mathbf{R}') can be calculated from the old inverse matrix by the relation

$$D_{jk}^{-1}(\mathbf{R}') = \begin{cases} D_{jk}^{-1}(\mathbf{R})/q, & k = i \\ D_{jk}^{-1}(\mathbf{R}) - \frac{1}{q} D_{ji}^{-1}(\mathbf{R}') \sum_{l=1}^N D_{lk}^{-1}(\mathbf{R}) D_{il}(\mathbf{R}'), & k \neq i \end{cases}, \quad (6.13)$$

where q is the ratio from Eq. (6.12).

The ratio for a multi-determinant wave function is

$$\frac{\sum_{i=1} c_i D_i(\mathbf{R}')}{\sum_{i=1} c_i D_i(\mathbf{R})} = \frac{\sum_{i=1} c_i q_i D_i(\mathbf{R})}{\sum_{i=1} c_i D_i(\mathbf{R})}, \quad (6.14)$$

where $q_i = D_i(\mathbf{R}')/D_i(\mathbf{R})$ from Eq. (6.12).

The derivatives are computed in a similar manner to the ratio in Eq. (6.12). To calculate the gradient with respect to particle i , expand the determinant in terms of its co-factors along row i and differentiate

$$\nabla_i D(\mathbf{R}) = \sum_{j=1}^N (\nabla_i D_{ij}) C_{ij}. \quad (6.15)$$

where $D_{ij}(\mathbf{R})$ is a function of \mathbf{r}_i and C_{ij} is a function of all of the other coordinates except

\mathbf{r}_i . This leads to the important relations:

$$\frac{\nabla_i D(\mathbf{R})}{D(\mathbf{R})} = \sum_{j=1}^N (\nabla_i D_{ij}(\mathbf{R})) D_{ji}^{-1}(\mathbf{R}). \quad (6.16)$$

and

$$\frac{\nabla_i^2 D(\mathbf{R})}{D(\mathbf{R})} = \sum_{j=1}^N (\nabla_i^2 D_{ij}(\mathbf{R})) D_{ji}^{-1}(\mathbf{R}). \quad (6.17)$$

6.3 Single-Particle Orbitals

The following sections describe the different types of basis functions used in this work to construct the single particle orbitals, including Slater Type Orbitals (STO)s and Gaussian Type Orbitals (GTO)s and Numerical Orbitals. For atoms these orbitals are commonly referred to as Atomic Orbitals (AO)s, while for molecules they are referred to as Molecular Orbitals (MO)s. The orbitals can be expressed in terms of a linear combination of basis functions centered at the sight of each atom.

The advantage of STOs over GTOs is that they have a form which is closer to an atomic-like orbital and therefore better satisfy the cusp conditions at the atomic nuclei. Unfortunately, STOs are more difficult to integrate than GTOs and analytic formulas do not exist.² This is the reason why GTOs are more commonly used in electronic structure codes. In general it is possible to fit a set of Gaussians to construct an STO; this approach is used to construct the basis for the hydrogen atom.

6.3.1 Gaussian Type Orbitals and the Gaussian 03 Basis

In the Gaussian03 package [11] the molecular orbitals are written as a linear combination of local orbitals

$$\psi_k(\mathbf{r}) = \sum_i \sum_j d_{ij} \phi_j(\mathbf{r} - \mathbf{R}_i), \quad (6.18)$$

²In Kohn-Sham DFT or Hartree-Fock methods the integrals are matrix elements of the Hamiltonian.

where the sum i runs over all the atoms and the sum j runs over all the local orbitals ϕ_j centered on atom i at position \mathbf{R}_i . The form of the local orbital is the sum of one or more primitive Gaussians multiplied by a polynomial and a real spherical harmonic

$$\phi_j(\mathbf{r}) = r^l S_{lm} \sum_{i=1}^N c_{i\mu} e^{-\alpha_{i\mu} f_\mu^2 r^2}, \quad (6.19)$$

where N is the number of primitive functions, *i.e.* the degree of contraction, the $\{c_{i\mu}\}$ are the contraction coefficients, the $\{\alpha_i\}$ are the exponents, f_μ is a scale factor common to all exponents and S_{lm} is a real spherical harmonic (see Appendix A.1.1). The factor f_μ allows for more variational freedom of the basis but for our applications it is usually set to one or can be absorbed into the $\{\alpha_i\}$.

It is convenient to factor out of the contraction coefficients a normalization for each primitive Gaussian: $c_i = N_l(\alpha_i) b_i$. The normalization factor $N_l(\alpha_i)$ is determined by the requirement

$$1 = \langle \phi_j | \phi_j \rangle, \quad (6.20)$$

since the S_{lm} are normalized this condition gives

$$1 = N_l^2(\alpha) \int_0^\infty dr r^{2(l+1)} e^{-2\alpha r^2}. \quad (6.21)$$

The above integral can be solved by using the formula

$$\int_0^\infty dr r^{2n} e^{-ar^2} = \frac{(2n-1)!!}{2^{n+1} a^n} \sqrt{\frac{\pi}{a}}, \quad (6.22)$$

which yields

$$N_l(\alpha) = 2^{l+1} \sqrt{\frac{2\alpha^{l+1}}{(2l+1)!!}} \left(\frac{2\alpha}{\pi} \right)^{\frac{1}{4}}. \quad (6.23)$$

Gaussian basis sets are often expressed in terms of *shells* [11]. A shell is defined as a set of basis functions $\{\phi\}$ with shared exponents $\{\alpha\}$. Examples of shells are s, p, sp, d, spd, \dots

An s -shell contains a single basis function with angular momentum $l = 0$; a p -shell contains three basis functions of p_x, p_y and p_z . Another example is the sp -type shell which is comprised of four basis functions s, p_x, p_y and p_z . For the sp shell the contraction coefficients must be the same for all functions of a given angular momentum. The basis sets used in this thesis are generally constructed of s, p, sp and d shells.

6.3.2 Slater Type Orbitals

The form of the Slater type orbital (STO) is given by

$$\phi_{nlm}(\mathbf{r}) = S_{lm} N_n(\xi) r^{n-1} e^{-\xi r}, \quad (6.24)$$

where n is the principal quantum number, ξ is an exponent and S_{lm} is a real spherical harmonic (see Appendix A.1.1). Following an identical procedure in Section 6.3.1, the normalization constant $N_n(\xi)$ is determined by the condition

$$1 = \langle \phi_{nlm} | \phi_{nlm} \rangle \quad (6.25)$$

$$= \int_0^\infty dr N_n(\xi)^2 r^{2n} e^{-2\xi r} \quad (6.26)$$

The integral can be solved by using the formula

$$\int_0^\infty dr r^{2n} e^{-2ar} = \frac{n\Gamma[2n]}{2^{2n} a^{2n+1}}, \quad (6.27)$$

yielding the result

$$N_n(\xi) = \sqrt{\frac{(2\xi)^{2n+1}}{(2n)!}}. \quad (6.28)$$

6.3.3 Numerical Orbitals

Numerical orbitals refer to any type of orbital on a grid. For a function on a *radial grid*, $f(r_i)$, with the endpoints r_i and r_f , it is possible to evaluate the function and its derivatives at any point r , not just at the grid points $\{r_i\}$ by use of an interpolation function. The interpolation function used in qmcPACK and the radial grid HF codes is the one-dimensional cubic spline discussed in Appendix A.3. For a discussion of the different types of radial grids see Appendix A.2.

Any analytic function, such as a GTO or an STO, can be discretized on a radial grid and interpolated. It can actually be faster to evaluate the functions numerically, by use of the cubic splines, than analytically. Many electronic structure codes such as the SIESTA DFT package [181] and our radial HF code use numerical orbitals. For a discussion the SIESTA Numerical Atomic Orbitals (NAO)s see Ref. [182]. Typically pseudopotentials are also evaluated on a radial grid. It is also possible to interpolate a three-dimensional grid function by use of a three-dimensional cubic spline. Many electronic structure codes use three-dimensional grid functions, and it is also possible to discretize the molecular orbitals from the Gaussian03 package [11] on a three-dimensional grid. For a discussion of three-dimensional cubic splines see Ref. [183].

6.4 The Electron-Electron Jastrow

The electron-electron Jastrow [179] correlates pairs of electrons

$$U_{ee}(\mathbf{R}) = \sum_{i < j} u_{ee}^{ij}(r_{ij}), \quad (6.29)$$

where u_{ee}^{ij} is a generic function that depends on the relative spins of the electron pair i and j , and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the relative electron-electron coordinate. For the sake of simplicity

all of the derivations of this section assume a spin-independent Jastrow function

$$u_{ee}^{ij}(r_{ij}) \Rightarrow u_{ee}(r_{ij}). \quad (6.30)$$

The electron-electron cusp condition (see Appendix A.7.2) must be satisfied by the electron-electron Jastrow. For a trial function of the form $\Psi = \exp(-U_{ee})$, the correct cusp condition is:

$$\left. \frac{1}{\Psi} \frac{\partial \Psi}{\partial r_{ij}} \right|_{r_{ij}=0} = - \left. \frac{\partial U_{ee}}{\partial r_{ij}} \right|_{r_{ij}=0} = \begin{cases} \frac{1}{2} & \text{for opposite spins} \\ \frac{1}{4} & \text{for parallel spins} \end{cases} \quad (6.31)$$

In practice usually only the opposite spin cusp condition is satisfied since like spin electrons are unlikely to approach each other by exchange interactions.

6.4.1 The Derivatives

To evaluate the gradient and Laplacian the following identities are useful:

$$\nabla_k(r_{ik}) = -\frac{\mathbf{r}_{ik}}{r_{ik}} \quad (6.32)$$

$$\nabla_k \cdot \left(\frac{\mathbf{r}_{ik}}{r_{ik}} \right) = -\frac{2}{r_{ik}} \quad (6.33)$$

$$r_{ik} = r_{ki}, \quad (6.34)$$

and the vector product rule

$$\nabla \cdot (f \mathbf{A}) = f(\nabla \cdot \mathbf{A}) + \mathbf{A} \cdot (\nabla f). \quad (6.35)$$

The first and second derivatives of the function u_{ee} are denoted by u'_{ee} and u''_{ee} , respectively.

The gradient

$$\begin{aligned}
\nabla_k U_{ee} &= \nabla_k \left[\sum_{i=1}^{k-1} u_{ee}(r_{ik}) + \sum_{i=k+1}^N u_{ee}(r_{ki}) \right] \\
&= - \sum_{i < k} u'_{ee}(r_{ik}) \frac{\mathbf{r}_{ik}}{r_{ik}} + \sum_{i > k} u'_{ee}(r_{ki}) \frac{\mathbf{r}_{ki}}{r_{ki}} \\
&= - \sum_{i \neq k} u'_{ee}(r_{ik}) \frac{\mathbf{r}_{ik}}{r_{ik}},
\end{aligned} \tag{6.36}$$

and the Laplacian

$$\begin{aligned}
\nabla_k^2 U_{ee} &= \sum_{i \neq k} \left[-u'_{ee}(r_{ik}) \left(\nabla_k \cdot \left(\frac{\mathbf{r}_{ik}}{r_{ik}} \right) \right) + \left(\frac{\mathbf{r}_{ik}}{r_{ik}} \right) \cdot (\nabla_k u'_{ee}(r_{ik})) \right] \\
&= \sum_{i \neq k} \left[\frac{2}{r_{ik}} u'_{ee}(r_{ik}) + u''_{ee}(r_{ik}) \right].
\end{aligned}$$

Summing over all k yields

$$\begin{aligned}
\nabla^2 U_{ee} &= \sum_k \nabla_k^2 U_{ee} \\
&= 2 \sum_{i < k} \left[\frac{2}{r_{ik}} u'_{ee}(r_{ik}) + u''_{ee}(r_{ik}) \right]
\end{aligned} \tag{6.37}$$

6.4.2 The Ratio

In QMC it is necessary to evaluate the ratio

$$r = \frac{\exp(U_{ee}(\mathbf{R}'))}{\exp(U_{ee}(\mathbf{R}))} = \exp(U_{ee}(\mathbf{R}') - U_{ee}(\mathbf{R})), \tag{6.38}$$

where $\mathbf{R}' = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}'_k, \dots, \mathbf{r}_N)$ differs from $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k, \dots, \mathbf{r}_N)$ only in the coordinates of the k th particle. The ratio can be shown to be equal to

$$r = \exp \left(\sum_{i \neq k} [u_{ee}(r_{ik'}) - u_{ee}(r_{ik})] \right), \quad (6.39)$$

where $r_{ik'} = |\mathbf{r}_i - \mathbf{r}'_k|$.

6.5 The Electron-Ion Jastrow

The electron-ion (single-electron) Jastrow correlates the electrons to the ions

$$U_{eI}(\mathbf{R}) = \sum_{i\alpha} u_{eI}^\alpha(r_{i\alpha}), \quad (6.40)$$

where u_{eI}^α is a generic function that depends on the ion species and $r_{i\alpha} = |\mathbf{r}_i - \mathbf{R}_\alpha|$ is the relative electron-ion coordinate. Note that the electron coordinates are denoted by \mathbf{r}_i , while the ion coordinates are \mathbf{R}_α . Besides electron-ion correlations, the single-electron Jastrow has been successfully used in quantum dot (QD) calculations where the coordinates \mathbf{R}_α would represent the centers of the QD potentials.

The electron-nuclear cusp condition (see Appendix A.7.1) can be satisfied by the single-electron Jastrow. The electron-nuclear cusp condition is discussed in great detail by Ma *et al.* [184]. In practice we are only interested in the cusp condition for the s ($l = 0$) electrons. For a trial function of the form $\Psi = \exp(-U_{eI}^\alpha)$, the correct cusp condition is:

$$\left. \frac{1}{\Psi} \frac{\partial \Psi}{\partial r_{i\alpha}} \right|_{r_{i\alpha}=0} = - \left. \frac{\partial U_{eI}}{\partial r_{i\alpha}} \right|_{r_{i\alpha}=0} = -Z_\alpha, \quad (6.41)$$

where Z_α is the charge of the ion. This is only necessary when the single-particle orbitals in the Slater determinant do not satisfy the correct nuclear cusp condition for the given ion species. If the bare Coulomb potential $V_\alpha(r) = -Z_\alpha/r$ of ion I is replaced by a pseudopo-

tential there is no cusp condition at \mathbf{R}_α .

The introduction of the single-electron Jastrow term serves to correct energetically unfavorable changes in the density caused by the electron-electron Jastrow correlation term [7, 31, 185]. In non-uniform systems the electron-electron Jastrow pushes electrons away from regions of high charge density into lower density regions. The introductions of the single-electron Jastrow alters the charge density without disturbing the pair-correlation function [31].

6.5.1 The Derivatives

To evaluate the gradient and Laplacian of U_{eI} , define the first and second derivatives of u_{eI}^α as $u'_{eI}{}^\alpha$ and $u''_{eI}{}^\alpha$, respectively.

$$\nabla_i U_{eI}(\mathbf{R}) = \sum_{\alpha} u'_{eI}{}^\alpha(r_{i\alpha}) \frac{\mathbf{r}_{i\alpha}}{r_{i\alpha}}. \quad (6.42)$$

and the Laplacian

$$\begin{aligned} \nabla_i^2 U_{eI}(\mathbf{R}) &= \sum_{\alpha} \left[u'_{eI}{}^\alpha(r_{i\alpha}) \left(\nabla_i \cdot \left(\frac{\mathbf{r}_{i\alpha}}{r_{i\alpha}} \right) \right) + \left(\frac{\mathbf{r}_{i\alpha}}{r_{i\alpha}} \right) \cdot (\nabla_i u'_{eI}{}^\alpha(r_{i\alpha})) \right] \\ &= \sum_{\alpha} \left[\frac{2}{r_{i\alpha}} u'_{eI}{}^\alpha(r_{i\alpha}) + u''_{eI}{}^\alpha(r_{i\alpha}) \right] \end{aligned} \quad (6.43)$$

6.5.2 The Ratio

In QMC it is necessary to evaluate the ratio

$$r = \frac{\exp(U_{eI}(\mathbf{R}'))}{\exp(U_{eI}(\mathbf{R}))} = \exp(U_{eI}(\mathbf{R}') - U_{eI}(\mathbf{R})), \quad (6.44)$$

where $\mathbf{R}' = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}'_k, \dots, \mathbf{r}_N)$ differs from $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k, \dots, \mathbf{r}_N)$ only in the coordinates of the k th particle. The ratio can be shown to be equal to

$$r = \exp \left(\sum_{\alpha} [u_{eI}^{\alpha}(r_{k'\alpha}) - u_{eI}^{\alpha}(r_{k\alpha})] \right), \quad (6.45)$$

where $r_{k'\alpha} = |\mathbf{r}'_k - \mathbf{R}_{\alpha}|$.

6.6 Examples of Jastrow Functions

The following sections contain different types of functions that are used for the one- and two-electron Jastrow correlation factors.

6.6.1 The Padé Jastrow

The generic form of the Padé-Jastrow [7] is given by

$$u(r) = \frac{a_1 r + a_2 r^2 + \dots}{1 + b_1 r + b_2 r^2 + \dots}, \quad (6.46)$$

where the most common form has only non-zero a_1 and b_1 :

$$u(r) = \frac{ar}{1 + br}, \quad (6.47)$$

with the derivatives

$$u'(r) = \frac{a}{(1 + br)^2} \quad (6.48)$$

$$u''(r) = \frac{-2ab}{(1 + br)^3}, \quad (6.49)$$

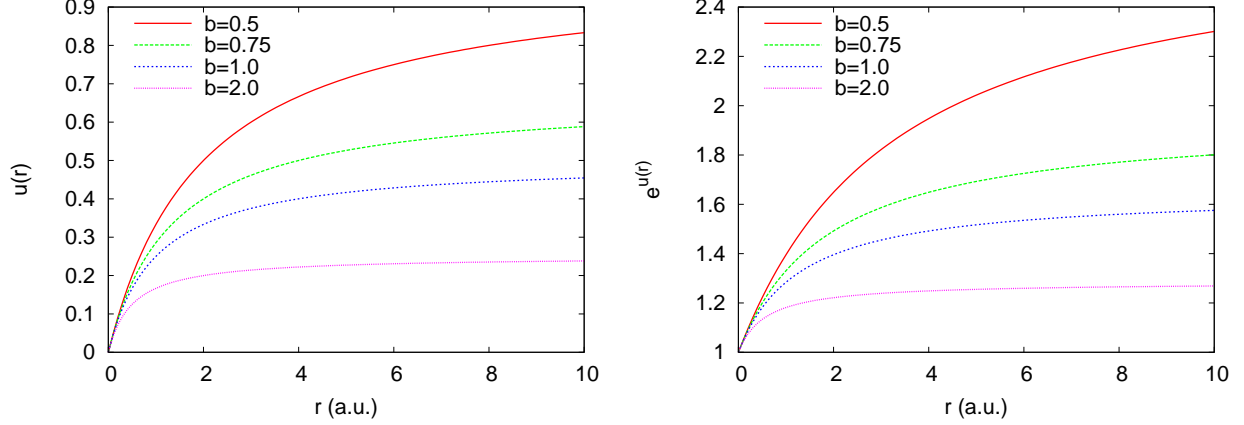


Figure 6.1: Padé-Jastrow $u(r)$ with $a = 1/2$ and several different values of b .

The asymptotic conditions are:

$$u(r)|_{r=0} = 0$$

$$u(r)|_{r \rightarrow \infty} = a/b$$

$$u'(r)|_{r=0} = a$$

To satisfy the electron-electron cusp conditions the parameter a is set to either $-1/2$ for opposite spin electrons or $-1/4$ for parallel spin electrons. For the case of a nuclear cusp condition $a = Z$, where Z is the charge of the nucleus. A plot of the Padé-Jastrow for several values of the parameter b is provided by Figure 6.1.

6.6.2 The Cusp-less Padé-Jastrow

The cusp-less Padé-Jastrow function is ideal for the cases where no cusp condition is required such as for harmonic potentials or pseudopotentials:

$$u(r) = \frac{a}{1 + br^2} \tag{6.50}$$

$$u'(r) = \frac{-2abr}{(1 + br^2)^2} \tag{6.51}$$

$$u''(r) = \frac{2ab(3br^2 - 1)}{(1 + br^2)^3}. \tag{6.52}$$

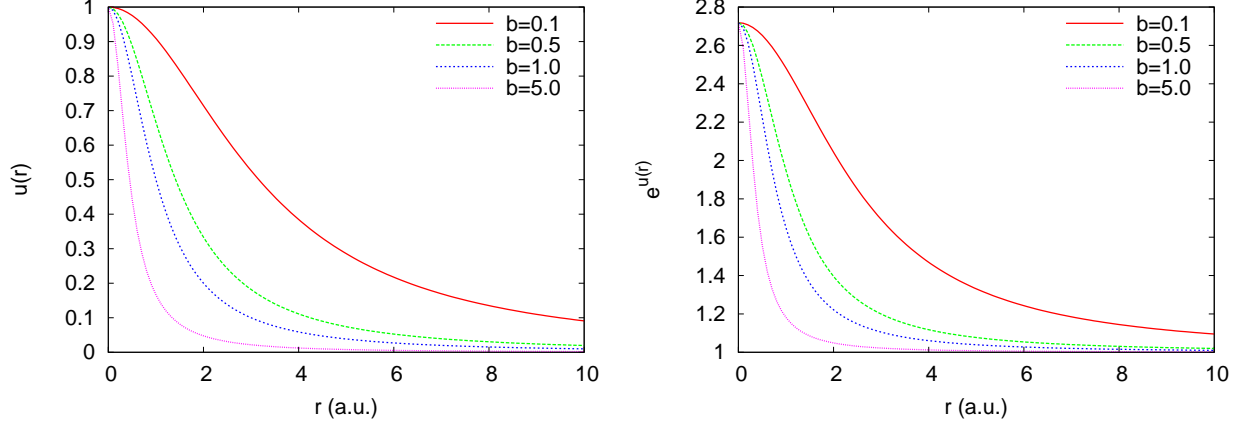


Figure 6.2: Cusp-Less Padé-Jastrow $u(r)$ with $a = 1.0$ and several values of b .

The asymptotic conditions are:

$$\begin{aligned}
 u(r)|_{r=0} &= a \\
 u(r)|_{r \rightarrow \infty} &= 0 \\
 u'(r)|_{r=0} &= 0.
 \end{aligned} \tag{6.53}$$

A plot of the cusp-less Padé-Jastrow for several values of the parameter b is provided by Figure 6.2.

6.6.3 The Long-Ranged Jastrow

This function is generally used for the electron-electron Jastrow for periodic systems [186]:

$$u(r) = \frac{a}{r} [1 - \exp(-r/b)] \tag{6.54}$$

$$u'(r) = \frac{-a}{r^2} \left[1 - \exp(-r/b) - \frac{r}{b} \exp(-r/b) \right] \tag{6.55}$$

$$u''(r) = \frac{2a}{r^3} \left[1 - \exp(-r/b) - \frac{r}{b} \exp(-r/b) - \frac{r^2}{2b^2} \exp(-r/b) \right]. \tag{6.56}$$

The asymptotic conditions are:

$$\begin{aligned} u(r)|_{r=0} &= a/b \\ u(r)|_{r \rightarrow \infty} &= a/r \\ u'(r)|_{r=0} &= -a/2b^2 \end{aligned} \tag{6.57}$$

The parameters a and b are not independent; to satisfy the opposite spin electron-electron cusp condition $b = \sqrt{a}$, while for parallel spin electrons $b = \sqrt{2a}$.

6.7 Evaluating the Local Energy

This section detailed how the local energy is evaluated for a trial wave function $\Psi_T(\mathbf{R})$. The local energy

$$E_L(\mathbf{R}) = \frac{H\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}, \tag{6.58}$$

is defined with respect to the N -electron Hamiltonian

$$H = -\frac{1}{2}\nabla^2 + V(\mathbf{R}), \tag{6.59}$$

where $\nabla^2 = \sum_{i=1}^N \nabla_i^2$ is the N -electron Laplacian and the potential $V(\mathbf{R})$ normally consists of an electron-electron interaction $V_{ee}(|\mathbf{r}_i - \mathbf{r}_j|)$ and an electron-ion interaction which is due to an external potential $V_{ext}(|\mathbf{r}_i - \mathbf{R}_\alpha|)$. For this section we assume that V_{ext} is a local potential, the evaluation of non-local potentials is presented in Section 7.6.

To evaluate the local energy it is convenient to differentiate the logarithm of the trial function

$$\begin{aligned} \nabla \exp(\ln \Psi) &= \nabla \ln \Psi \exp(\ln \Psi) \\ &= [\nabla(\ln \Psi)]\Psi_T \\ \nabla^2 \exp(\ln \Psi) &= [\nabla^2(\ln \Psi) + (\nabla(\ln \Psi))^2]\Psi_T. \end{aligned} \tag{6.60}$$

The kinetic energy (KE) part of the local energy

$$KE = -\frac{1}{2}\Psi_T^{-1}\nabla^2\Psi_T = -\frac{1}{2}\left[\nabla^2(\ln\Psi) + (\nabla(\ln\Psi))^2\right]. \quad (6.61)$$

can be expressed in terms of the quantities

$$T = -\frac{1}{4}\nabla^2(\ln\Psi) = -\frac{1}{4}\frac{\nabla^2\Psi_T}{\Psi_T} + \frac{1}{4}\left(\frac{\nabla\Psi_T}{\Psi_T}\right)^2 \quad (6.62)$$

and

$$\mathbf{F} = \frac{1}{\sqrt{2}}\nabla(\ln\Psi) = \frac{1}{\sqrt{2}}\frac{\nabla\Psi_T}{\Psi_T}, \quad (6.63)$$

where the relation is given by

$$KE = 2T - |\mathbf{F}|^2. \quad (6.64)$$

The $3N$ -dimensional vector \mathbf{F} is closely related to the drift velocity $\mathbf{V}_D = \Psi_T^{-1}\nabla\Psi_T$ at the position \mathbf{R} . By use of the Green's Theorem and integration by parts it is possible to derive the relation:

$$\langle KE \rangle = \langle T \rangle = \langle |\mathbf{F}|^2 \rangle, \quad (6.65)$$

where the averages $\langle \rangle$ are computed with respect to Ψ_T^2 . The relationship in Eq. (6.65) can be used to check if Ψ_T^2 is sampled correctly and if Ψ_T and its derivatives are being correctly evaluated, see Figure 6.3 for an example.

The gradient and Laplacian of $\ln\Psi$, when $\Psi_T(\mathbf{R}) = D(\mathbf{R})\exp(-U(\mathbf{R}))$ is Slater-Jastrow type, are given by

$$\nabla(\ln\Psi) = -\nabla U + \frac{\nabla D}{D} \quad (6.66)$$

and

$$\nabla^2(\ln\Psi) = -\nabla^2 U + \frac{\nabla^2 D}{D} - \left(\frac{\nabla D}{D}\right)^2, \quad (6.67)$$

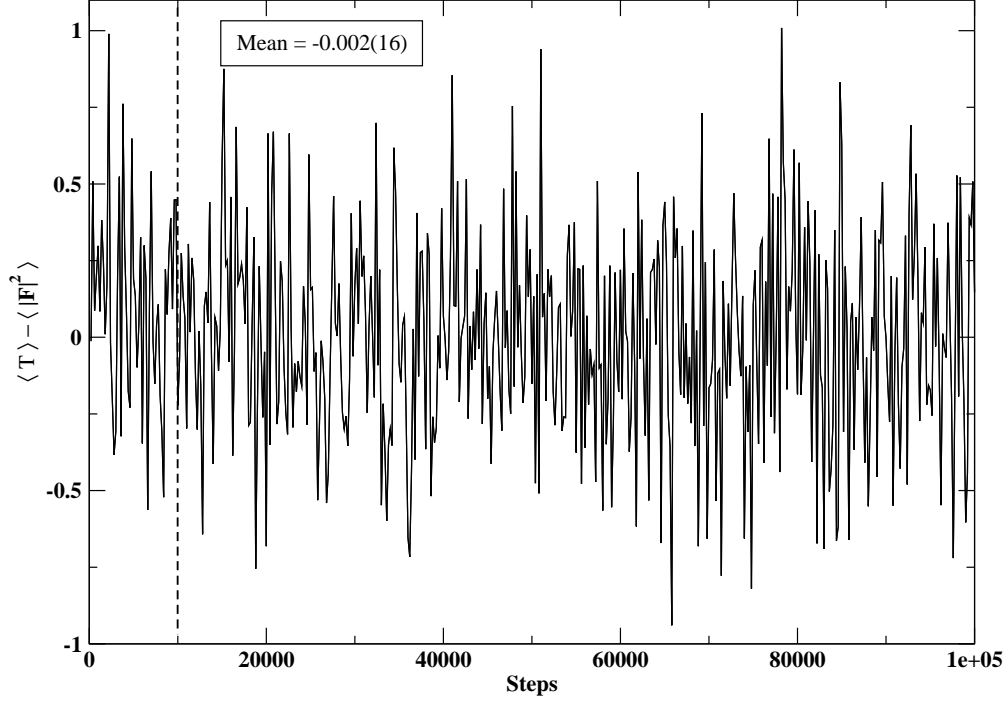


Figure 6.3: The VMC estimator of $\langle T \rangle - \langle |\mathbf{F}|^2 \rangle$ for the Li_2 molecule.

where ∇ is the $3N$ -dimensional gradient. The final expression for the local energy is:

$$E_L = -\frac{1}{2} \left[\frac{\nabla^2 D}{D} - \nabla^2 U + (\nabla U)^2 - 2 \nabla U \cdot \frac{\nabla D}{D} \right] + V. \quad (6.68)$$

Expressions for evaluating the derivatives of the determinant and Jastrow parts (electron-electron and electron-ion) of Ψ_T can be found in Sections 6.2, 6.4 and 6.5.

For a multi-determinant trial wave function the gradient and Laplacian are:

$$\nabla \ln \Psi = -\nabla U + \frac{(\sum_{n=1}^M c_n \nabla D_n)}{\sum_{n=1}^M c_n D_n}, \quad (6.69)$$

and

$$\nabla^2 \ln \Psi = -\nabla^2 U + \frac{\sum_{n=1}^M c_n \nabla^2 D_n}{\sum_{n=1}^M c_n D_n} - \left(\frac{\sum_{n=1}^M c_n \nabla D_n}{\sum_{n=1}^M c_n D_n} \right)^2. \quad (6.70)$$

6.8 Optimization of the Trial Wave Function

6.8.1 The Cost Function and Correlated Sampling

The Slater-Jastrow trial function Ψ_T is parameterized by the set of variational parameters α . The goal of optimization is to find the set α that minimizes a cost function. A natural choice for the cost function is the expectation value of the local energy

$$E_V(\alpha) = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha) E_L(\mathbf{R}; \alpha)}{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha)}, \quad (6.71)$$

where $E_L(\alpha) = \Psi_T^{-1}(\alpha) H \Psi_T(\alpha)$. Minimizing the variance of the local energy

$$\sigma_E^2(\alpha) = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha) [E_L(\mathbf{R}; \alpha) - E_V(\alpha)]^2}{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha)} \quad (6.72)$$

is desirable due the variance minimization principle

$$\sigma^2[H] = \langle H^2 \rangle_{\Psi_T^2} - \langle H \rangle_{\Psi_T^2}^2 \geq 0 \quad (6.73)$$

which states that the absolute minimum value of the variance is zero. Since the absolute minimum of Eq. (6.72) occurs when Ψ_T is *any* eigenfunction of H , variance minimization is suitable for excited states as well. In practice excited state wave functions are usually not optimized. It is also possible to optimize a linear combination of E_V and σ_E^2 , but it should be noted that the energy and variance have different units.

A variant of Eq. (6.72) involves minimizing the second moment of the local energy with respect to a fixed reference energy E_R :

$$\sigma_{E_R}^2 = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha) [E_L(\mathbf{R}; \alpha) - E_R]^2}{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha)}, \quad (6.74)$$

where the constant E_R should be close to the energy of the state being sought.³

In general it is possible to minimize any function of the form

$$\frac{\int d\mathbf{R} P(\mathbf{R}; \boldsymbol{\alpha}) f(E_L(\mathbf{R}; \boldsymbol{\alpha}) - E_R)}{\int d\mathbf{R} P(\mathbf{R}; \boldsymbol{\alpha})}, \quad (6.75)$$

where P is a probability distribution and f is a non-negative function such that $f(0) = 0$ [187]. qmcPACK can optimize the class of functions

$$f(E_L(\mathbf{R}; \boldsymbol{\alpha}) - E_R) = |E_L(\mathbf{R}; \boldsymbol{\alpha}) - E_R|^n, \quad (6.76)$$

for any integer $n \geq 1$ over the distribution $P = \Psi_T^2(\boldsymbol{\alpha})$.

The cost function is computed as an average over a finite set of configurations generated by VMC. A new set of configurations is required each time $\boldsymbol{\alpha}$ is changed in the optimization process; this is necessary to correctly sample $\Psi_T^2(\boldsymbol{\alpha})$. By using the correlated sampling approach (Section 4.8), only a single set of configurations from the distribution $\Psi_T(\boldsymbol{\alpha}_0)^2$ is required, where $\boldsymbol{\alpha}_0$ is the initial parameter set [155]. The cost function in Eq. (6.74) is evaluated as:

$$\sigma_{E_R}^2 = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \boldsymbol{\alpha}) w(\mathbf{R}; \boldsymbol{\alpha}) (E_L(\mathbf{R}; \boldsymbol{\alpha}) - E_R)^2}{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \boldsymbol{\alpha}) w(\mathbf{R}; \boldsymbol{\alpha})}, \quad (6.77)$$

where the weights are

$$w(\mathbf{R}; \boldsymbol{\alpha}) = \frac{\Psi_T(\mathbf{R}; \boldsymbol{\alpha})}{\Psi_T(\mathbf{R}; \boldsymbol{\alpha}_0)}, \quad (6.78)$$

and $\boldsymbol{\alpha}$ is the current parameter set. The discrete approximation to Eq. (6.77) is given by

$$\sigma_{E_R}^2 \approx \frac{\sum_{i=1}^M w(\mathbf{R}_i; \boldsymbol{\alpha}) (E_L(\mathbf{R}_i; \boldsymbol{\alpha}) - E_R)}{\sum_{i=1}^M w(\mathbf{R}_i; \boldsymbol{\alpha})}, \quad (6.79)$$

where the number of configurations M is usually of the order $10^3 - 10^5$.

An improved method for optimization when calculating extrapolated estimators (Sec-

³Minimizing Eq. (6.74) is equivalent to minimizing a linear combination of E_V and σ_E^2 , for a proof see pg. of Ref. [7].

tion 4.10.4) is to maximize the overlap between the trial wave function Ψ_T and the projected wave function $\Phi_0(\beta \rightarrow \infty)$ within DMC [188]

$$\frac{\int d\mathbf{R} \Psi_T(\mathbf{R}; \boldsymbol{\alpha}) \Phi_0(\mathbf{R})}{\left[\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \boldsymbol{\alpha}) \right]^{1/2}}. \quad (6.80)$$

The actual minimization of the cost function can be carried out by using standard optimization techniques. The most common methods are conjugate gradient, steepest descent and simulated annealing. The default method of qmcPACK is the conjugate gradient algorithm from Numerical Recipes [189]. Algorithms that require analytic evaluation of the cost function and its derivatives are avoided. The parameters $\boldsymbol{\alpha}$ are varied until the convergence condition on the cost function is satisfied. It may be necessary to introduce constraints on $\boldsymbol{\alpha}$ to prevent divergences in trial function $\Psi_T(\boldsymbol{\alpha})$ and the local energy $E_L(\boldsymbol{\alpha})$.

6.8.2 Optimizing the Jastrow

Usually only the Jastrow part of the Slater-Jastrow trial wave function is parameterized:

$$\Psi_T(\mathbf{R}; \boldsymbol{\alpha}) = D(\mathbf{R}) \exp[-U(\mathbf{R}; \boldsymbol{\alpha})]. \quad (6.81)$$

When this is the case the weights, Eq. (6.78), are determined by the ratio of the Jastrow factors:

$$w(\mathbf{R}; \boldsymbol{\alpha}) = \exp[-U(\mathbf{R}; \boldsymbol{\alpha}) + U(\mathbf{R}; \boldsymbol{\alpha}_0)]. \quad (6.82)$$

Only the terms in the local energy, Eq. (6.68), involving the Jastrow need to be re-evaluated during the optimization. For an example of the optimization of the Jastrow for the Ge atom, see Figure 6.4.

In general it is possible to optimize the single particle orbitals in the Slater determinant if they are parameterized. This is idea behind the correlated geminal approach used by Casula *et al.* [190]. Also it is possible to optimize the coefficients of the Slater determinants for

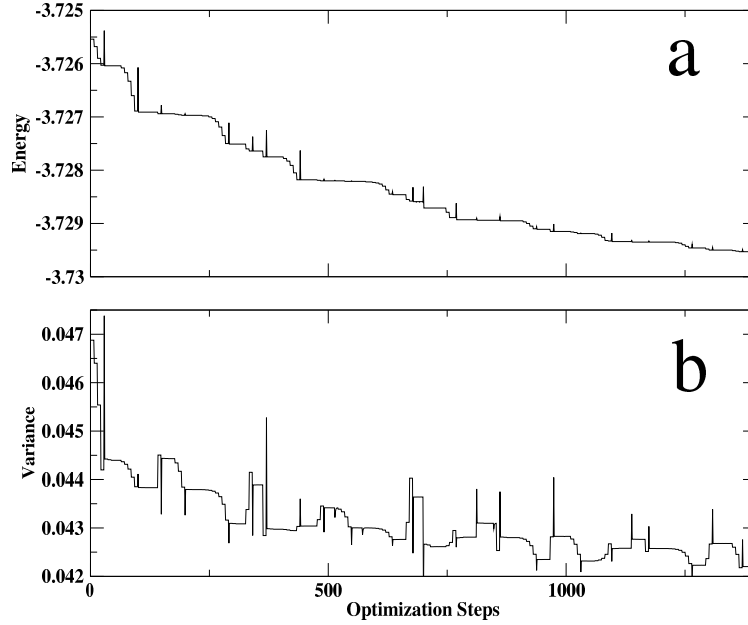


Figure 6.4: Optimization of the ground-state of the Ge atom (3P) using a cost function which is a linear combination of Eqs. (6.71) and (6.72). The optimizable parameters are in the one- and two-electron Jastrow functions; for details regarding the trial wave function see Section 8.2. a) Plots the energy for each step of the optimization and b) plots the variance of the local energy. The optimization algorithm used in this example was conjugate gradient. All units in a.u.

multi-determinant trial wave functions.

6.8.3 Optimization in the Presence of a Non-Local Potential

For a local potential, the potential energy (PE) is independent of the trial wave function and depends only on the positions of the particles. In contrast, for non-local potentials, the potential energy depends explicitly on the trial wave function, see Eq. (7.20). Typically the non-local potential is split into a local part common to all angular momentum channels and an angular-momentum dependent correction to the local part. Since the non-local corrections only constitute a small part of the total energy, it is not necessary to recalculate the potential part of the local energy during the optimization. This has the effect of saving a considerable amount of computation time [191].

6.8.4 Numerical Instability

The optimization procedure outlined above is often successful, but can be numerically unstable. The optimization can reach a minimum for a parameter set α that gives worse results for subsequent QMC calculations than the original parameter set α_0 , in terms of both the local energy and the variance of the local energy. The origin of this problem can be traced to the presence of walkers acquiring a large weight, usually only a single walker, which ruin the minimization [187]. In principle this problem could be remedied by increasing the number of walkers, but this is often impractical and the number required can be prohibitively large. A method for dealing with the presence of *bad* walkers in the finite walker set is to determine the effective number of walkers

$$N_{eff} = \frac{\left[\sum_{i=1}^N w(\mathbf{R}_i) \right]^2}{\sum_{i=1}^N w(\mathbf{R}_i)}, \quad (6.83)$$

where the total number of walkers is N . If all the weights are close to unity the effective number of walkers is close to N , $N_{eff} \approx N$: in the extreme case where all the weight is concentrated in one walker, the effective weight is close to unity, $N_{eff} \approx 1$. It is advisable to terminate the optimization if N_{eff} is less than a cutoff, which is usually around 80% of the total number of walkers. The problem can usually be traced to using a poor guess for α_0 or an insufficient number of walkers.

Numerical instability in the weights is especially prevalent for systems containing large numbers of electrons such as molecular clusters and bulk materials. We found this to be directly relevant to the Hydrogen passivated Si and Ge nanoclusters larger than Si_2H_6 or Ge_2H_6 . Fortunately, in the variance minimization scheme it is possible to alter the weights while not affecting the positions of the minima. Constraining the weights of the variance to be unity [192, 31, 51] ($w = 1$) yields

$$\sigma_E^2(\alpha) = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha_0) [E_L(\mathbf{R}; \alpha) - E_C(\alpha)]^2}{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha_0)} \quad (6.84)$$

where the un-weighted energy is

$$E_C(\alpha) = \frac{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha_0) E_L(\mathbf{R}; \alpha)}{\int d\mathbf{R} \Psi_T^2(\mathbf{R}; \alpha_0)}. \quad (6.85)$$

The absolute minima of Eq. (6.84) at zero occurs if $\Psi_T(\alpha)$ is an exact eigenfunction of H , because for an exact eigenfunction $E_L = E_C$. Therefore, the absolute minima of Eqs. (6.84) and (6.72) are at the same positions, but Eq. (6.84) is advantageous since it has a lower variance. The function E_C does not share this property, the ground-state of H is not guaranteed to be the minimum value of E_C . For an example of constrained weight optimization of the $\text{Ge}_{29}\text{H}_{36}$ nanocluster, see Figure 6.5.

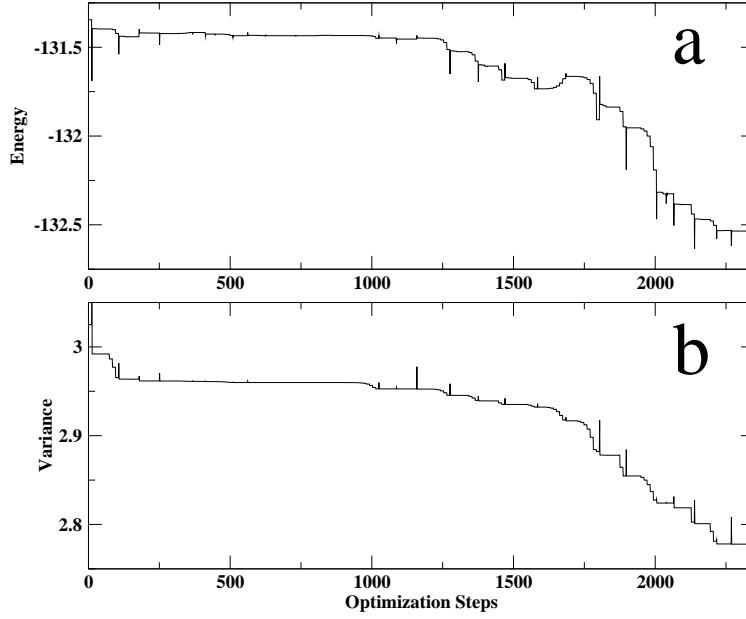


Figure 6.5: Optimization with constrained weights of the ground-state of the $\text{Ge}_{29}\text{H}_{36}$ nanocluster using the variance as the cost function. The optimizable parameters are in the one- and two-electron Jastrow functions; for details regarding the trial wave function see Section 8.2. a) Plots the energy for each step of the optimization and b) plots the variance of the local energy.

6.8.5 Outline of the Algorithm

This is the outline of the algorithm to optimize a trial wave function $\Psi_T(\boldsymbol{\alpha})$, see Algorithm 2. As a first step it is always necessary to generate a set of walkers from VMC to sample the initial distribution $\Psi_T(\boldsymbol{\alpha}_0)^2$.

It is always advisable to first perform an optimization on a simpler problem before proceeding to the more difficult problem under consideration. For example, to optimize a small molecule such as GeH_4 it is recommended to first optimize the trial wave functions for the constituent Ge and H atoms. The atomic results can be used as a good initial guess for trial wave function for GeH_4 . To optimize a larger Ge structure, such as Ge_5H_{12} , it is recommended to use as an initial guess the results from the optimization of GeH_4 .

Algorithm 2 Optimization of the trial wave function $\Psi_T(\boldsymbol{\alpha})$. The value of x used to determine the cutoff is usually chosen to be 0.8.

Generate a set of N configurations from the distribution $\Psi_T(\boldsymbol{\alpha}_0)^2$.

repeat

Evaluate the Cost function (For example Eq. (6.77)). The weights are either evaluated by Eq. (6.78) or are set to 1.0.

Calculate the effective number of walkers N_{eff} , Eq. (6.83).

if $N_{eff} < xN$ **then**

Stop the optimization.

end if

Adjust the parameters using algorithm of choice: $\boldsymbol{\alpha}_i \longrightarrow \boldsymbol{\alpha}_{i+1}$

Update the trial function: $\Psi_T(\boldsymbol{\alpha}_i) \longrightarrow \Psi_T(\boldsymbol{\alpha}_{i+1})$.

until Convergence criteria of the algorithm is satisfied.

If $\boldsymbol{\alpha}_i$ differs significantly from $\boldsymbol{\alpha}_0$ it is advisable to rerun VMC for $\Psi_T(\boldsymbol{\alpha}_i)$.

Chapter 7

Core-Valence Partitioning

7.1 Introduction

Core-valence partitioning schemes allow electronic structure calculations only on valence electrons by quantifying the effects the core electrons have on the valence electrons. Such schemes are successful because the outer valence electrons determine most of the chemical properties of atoms and molecules while the core electrons are relatively inert. The most common implementation of core-valence partitioning is by a pseudopotential, or effective core potential. Pseudopotentials are desirable because they allow the elimination of the core electrons, hence reducing the effective atomic number, and feature valence orbitals that are smooth and node-less. In the context of many-body methods such as Quantum Monte Carlo (QMC), pseudopotentials enable calculations that would otherwise be unfeasible. The scaling of QMC algorithms with respect to the atomic number Z is highly nonlinear because shorter time steps and longer runs are required for the higher energy core states. Estimates vary [33, 7] from $Z^{5.5}$ to $Z^{6.5}$ for all-electron calculations. Typically pseudopotentials are constructed within single-body methods such as density functional theory (DFT) in the local density approximation (LDA) or Hartree-Fock (HF), which means that the core-valence interactions are treated at the same levels. While this approach is sufficient for most elements in the periodic table, where core-valence interactions are negligible, it is unsatisfactory for post-transition elements such as Ge which has shallow and easily polarizable $3d$ core.

The salient issue is that the accuracy of QMC applied to heavy atoms such as Ge ($Z = 32$) is limited by the treatment of the core-valence interactions. No matter how accurate the

QMC calculations for the valence electrons, the accuracy is limited by the pseudopotential. Therefore, an improved method of core-valence partitioning is necessary for applications of QMC to elements such as Ge.

Shirley and Martin [43] have implemented a many-body version of core-valence partitioning based on the core polarization potential [40, 41, 42] (CPP) approach. Their parameterized CPPs are fit to quasiparticle self-energies, which include many-body effects such as core-relaxation and polarization and core-valence correlation. This is significant for alkali metals and post-transition elements such as Zn, Ga and Ge with shallow cores, core-valence correlations cannot be ignored. Specifically, the CPPs are purely *ab initio* and correct a HF treatment of core-valence interactions [43].

The effect of core-valence and intershell correlation is especially pronounced on the band structure of Ge. Shirley, Zhu and Louie [9] found significant improvement in the theoretical quasiparticle band energies of Ge when core-valence correlation was included via a CPP, see Table 7.1. Their CPP based results predicted the correct ordering of the closely spaced conduction band states, including the $\Gamma - L$ indirect band gap. The results based on an LDA treatment of the core-valence interactions were not nearly as accurate, despite the fact that the valence-valence interactions were treated at the same level of accuracy.¹

For the Ge_2H_6 molecule, Nicklass and Stoll (NS) [195] determined that core-valence correlation is a significant effect in determining the geometry, especially the Ge-Ge bond length. NS were able to account for the influence of intershell correlation by including the $3d$ orbitals in the active space of all-electron CCSD calculations. These results were corroborated by pseudopotential ones, where the core-valence correlation effects were modeled by a CPP [196, 197].

¹Ku and Eguiluz [193], in a paper reporting all-electron quasiparticle calculations for Ge, have attacked the work of Shirley *et al.* [9] due to their treatment of the core electrons. The authors of the earlier work have replied with criticism that Ku and Eguiluz did not treat the valence-valence correlations well enough [194].

	Quasiparticle Results			
	Core-Valence Interactions:			Expt.
	LDA	LDA	HF+CPP	
$\Gamma_{8v} \rightarrow \Gamma_{7c}$	-0.26	0.53	0.85	0.89
$\Gamma_{8v} \rightarrow X_{5c}$	0.55	1.28	1.09	1.10 [†]
$\Gamma_{8v} \rightarrow L_{6c}$	-0.05	0.70	0.73	0.744
$\Gamma_{8v} \rightarrow X_{5c}$	0.60	0.58	0.36	0.36

Table 7.1: Comparison of band energy differences for Ge from LDA and quasiparticle results to experiment from Ref [9]. The quasiparticle results feature two different treatments of the core-valence interactions: LDA and a CPP based treatment. The fundamental gaps are in bold. All energies in eV. Experimental results from Ref. [10] unless otherwise noted.

[†] From Ref. [9] and references therein.

7.2 Introduction to Pseudopotentials

It is possible to remove the core electrons by introducing a pseudopotential (PP), or an effective core potential according to quantum chemists. The use of pseudopotentials is justified since chemical properties, such as bond strengths, ionization energies and molecular geometries are largely determined by the valence electrons: the core electrons are spectator electrons and the core states retain their atom-like character. The disparity between the energy scales of the core electrons and valence electrons provides further justification for this decoupling [198, 40].

The effect of using a pseudopotential for a given atomic species is to reduce any electronic structure calculation to only the valence electrons. For the case of Ge, this reduces the calculation from 32 electrons to only 4 valence electrons:

$$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^6(3d)^{10}(4s)^2(4p)^2 \rightarrow (4s)^2(4p)^2. \quad (7.1)$$

Before going any further it is important to discuss an important aspect of pseudopotentials known as *transferability*. For a pseudopotential to be truly useful, it must not only mimic the atoms full-atomic valence properties for a given atomic reference configuration, but should also be transferable to different chemical environments, such as excited atomic

configurations, or more importantly molecules and solids.

Our interest is mainly confined to a class of *ab-initio* pseudopotentials known as *norm-conserving* pseudopotentials (NCPP)s. NCPPs are constructed to be equal to the all-electron (full-atomic) potential outside the core region, but are shallow and are typically *angular momentum dependent* inside the core region. The angular momentum dependence allows the pseudopotential to bind pseudo-valence orbitals with eigenvalues equal or nearly identical to the full-atomic eigenvalues for each angular momentum channel of interest. In addition, NCPPs are constructed to have nearly identical scattering properties as the full-atomic potential over a reasonably wide energy range for each angular momentum channel of interest. As we will see these properties make NCPPs highly transferable.

Due to the angular momentum dependence, the pseudopotential takes the form of a non-local operator; the *semi-local* form is given by

$$\hat{V}^{PP} = \sum_{lm} |Y_{lm}\rangle V_l(r) \langle Y_{lm}|, \quad (7.2)$$

where Y_{lm} is a spherical harmonic and $V_l(r)$ is a radial function. The pseudopotential is termed as being semi-local because it is only non-local in the angular variables (θ, ϕ) , but it is local in the radial variable.

7.2.1 Norm-Conserving Pseudopotentials

Norm-conserving pseudopotentials feature valence pseudo-orbitals that are smooth and nodeless, and when normalized become identical to the all-electron orbitals beyond a core radius r_c . For an excellent review of NCPPs see Refs. [199, 200, 201, 202, 203, 29, 204].

Hamann, Schlüter and Chiang (HSC) [200] listed a set of criteria for good NCPPs. (It should be noted that the principle quantum number n is omitted; the all-electron orbital is designated ψ^{AE} and the pseudo orbital is ψ^{PP} . Also, it is assumed that the valence pseudo orbitals satisfy the usual orthonormality conditions: $\psi_i^{PP} | \psi_j^{PP} \rangle = \delta_{ij}$.)

The HSC criteria:

A. The all-electron and pseudo eigenvalues agree

$$\epsilon_l^{AE} = \epsilon_l^{PP}. \quad (7.3)$$

B. The normalized all-electron and the normalized pseudo orbitals agree beyond the radius r_c , which is known as the core-radius

$$\psi_l^{AE}(r) = \psi_l^{PP}(r) \quad \text{for } r > r_c. \quad (7.4)$$

C. The partial-norm integral from 0 to r_c , *i.e.* the integrated charge density, of the all-electron and pseudo charge densities agree beyond r_c .

$$\int_0^{r_c} dr r^2 |\psi_l^{AE}(r)|^2 = \int_0^{r_c} dr r^2 |\psi_l^{PP}(r)|^2. \quad (7.5)$$

This is the norm conservation condition.

D. The logarithmic derivatives and the first energy derivative of the logarithmic derivative of the all-electron and pseudo wave functions agree at r_c

$$\left. \frac{d}{dr} \ln \psi_l^{PP} \right|_{r=r_c} = \left. \frac{d}{dr} \ln \psi_l^{AE} \right|_{r=r_c} \quad (7.6)$$

$$\left. \frac{\partial}{\partial \epsilon} \frac{d}{dr} \ln \psi_l^{PP} \right|_{r=r_c} = \left. \frac{\partial}{\partial \epsilon} \frac{d}{dr} \ln \psi_l^{AE} \right|_{r=r_c} \quad (7.7)$$

The radial logarithmic derivative is related to the scattering phase shift.

From properties 1 and 2 it follows that the NCPP equals the AE potential outside the core region because the potential is uniquely determined by the wave function and the energy ϵ (not necessarily an eigenvalue). Property 3 guarantees, through Gauss theorem, that the

electrostatic potential produced outside of r_c is identical for the all-electron and pseudo charge densities. This property holds despite the fact that ψ_l^{PP} and ψ_l^{AE} differ inside the core region. Property 4 guarantees that the PP and AE potential have the same scattering phase shifts (at r_c) to linear order in energy around ϵ_l . This is important because changing the chemical environment shifts the eigenvalues away from the atomic energy levels from the reference configuration.

Properties 3 and 4 are crucial for the pseudopotential to have the optimum transferability and are related by the important identity

$$-2\pi \left[(r\psi)^2 \frac{\partial}{\partial \epsilon} \frac{d}{dr} \ln \psi \right]_{r=r_c} = 4\pi \int_0^{r_c} dr r^2 \psi^2, \quad (7.8)$$

where ψ is a solution to the radial Schrodinger equation with the energy ϵ (not necessarily an eigenvalue). Eq. (7.8) implies that if ψ_l^{PP} and ψ_l^{AE} have the same magnitude at r_c and the same integrated charge density inside a sphere of radius r_c , then the first energy derivative of the logarithmic derivative are the same for ψ_l^{PP} and ψ_l^{AE} . For a proof of the identity in Eq. (7.8), see Appendix A.8. In addition Shirley *et al.* [205] observed that for NCPPs the total charge inside r_c divided by $\psi(r_c)^2$,

$$\frac{1}{\psi(r_c)^2} \int_0^{r_c} dr r^2 \psi(r)^2 \quad (7.9)$$

is the same for ψ^{PP} and ψ^{AE} .

7.3 Generating Pseudopotentials

Pseudopotentials are normally constructed to reproduce the all-electron results of the valence electrons of an atomic reference configuration: such as eigenvalues, partial-norms and the radial logarithmic derivatives. Typically the reference configuration is the ground state of the atom, although it is possible to use an excited state configuration or even an ionized

state. There are many different procedures for generating norm-conserving pseudopotentials. One of the most popular and widely used schemes is by Troullier and Martins [206].

One of the most difficult parts of generating any pseudopotential is in choosing the matching radii r_c for each of the angular momentum channels. For NCPP, r_c must be larger than outermost node of the orbital - this represents a lower bound for r_c . Generally r_c is chosen to be near the outermost peak, the larger the r_c the *softer* the pseudopotential, but this comes at the price that the pseudopotential is less transferable to different chemical environments [204, 199]. The softness of a pseudopotential refers to how many basis functions or plane waves are required to describe the pseudopotential.

7.4 The Germanium Pseudopotential

For Ge we use a relativistic Hartree-Fock (HF) pseudopotential from Ref. [12] to replace the core electrons, including the $3d$ -shell. This pseudopotential was used for all of our Gaussian 03 and QMC calculations unless otherwise specified. We have also created a local pseudopotential for Ge which is discussed in Chapter A.11. The pseudopotential reduces the calculation for each Ge atom from 32 electrons to only four valence electrons: $4s^2 4p^2$. The main reason for this choice of pseudopotential is that the SM CPP was developed as an addendum to HF based treatments of the core-valence interactions. In addition, there is also a growing body of work that shows that HF pseudopotentials perform better than DFT pseudopotentials in QMC [207], but it should be stated that this work is not conclusive.

The details in the generation of the pseudopotential can be found in Refs. [12, 208] and are briefly summarized in Ref [209]. The pseudopotential was generated in the Troullier-Martins [206] scheme. For use in non-relativistic calculations it is necessary to perform a j weighted averaging. This is because the valence electrons are specified by the quantum number κ ; the relationship between κ, j and l is summarized in Eq. (2.114). Therefore there are different pseudopotentials V_κ with the same l values. Taking the average of the pseudopotentials with

the same l quantum numbers weighted by the different j degeneracies leads to:

$$V_l = \frac{1}{2l+1} [lV_{\kappa=l} + (l+1)V_{\kappa=-l-1}]. \quad (7.10)$$

The spin-orbit coupling can be included separately.

It was necessary to remove the “extreme non-locality” of the pseudopotential by slightly relaxing the norm-conservation condition. The extreme non-locality refers to the long-range behavior of the pseudopotential: the pseudopotential is non-local over all space and deviates from the Coulomb $-Z/r$ potential in the $r \rightarrow \infty$ limit. This is a consequence of the non-locality of the exchange interaction in Hartree-Fock (Dirac-Fock). Figure 7.1 plots the Ge pseudopotential.

7.5 Pseudopotentials in Gaussian

The form of the pseudopotential used in the Gaussian 03 [11] package

$$\hat{V}^{PP} = V_{loc} + \hat{V}_{NL}, \quad (7.11)$$

where V_{loc} is the local component and \hat{V}_{NL} is the non-local correction to V_{loc} . The pseudopotential is represented in terms of a Gaussian expansion by:

$$V_{loc}(r) = \sum_{k=1}^M r^{n_k} c_k e^{-\alpha_k r^2}, \quad (7.12)$$

and

$$\hat{V}_{NL} = \sum_{l=1}^{l_{max}} \left[\sum_{k=1}^{M_l} r^{n_k} c_k e^{-\alpha_k r^2} \right] \hat{P}_l, \quad (7.13)$$

where the c_k are expansion coefficients and the operator $\hat{P}_l = |Y_{lm}\rangle\langle Y_{lm}|$ is the angular momentum projector operator.

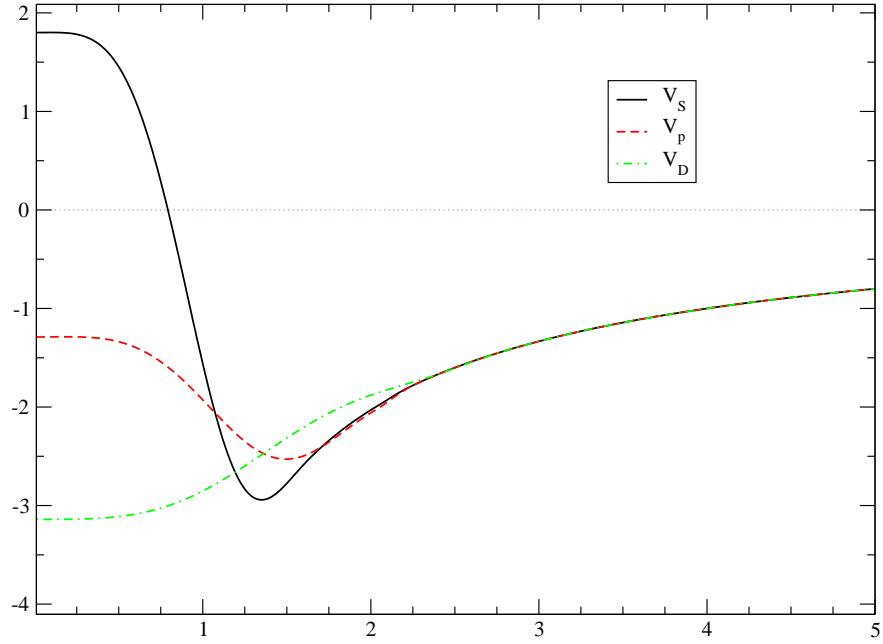
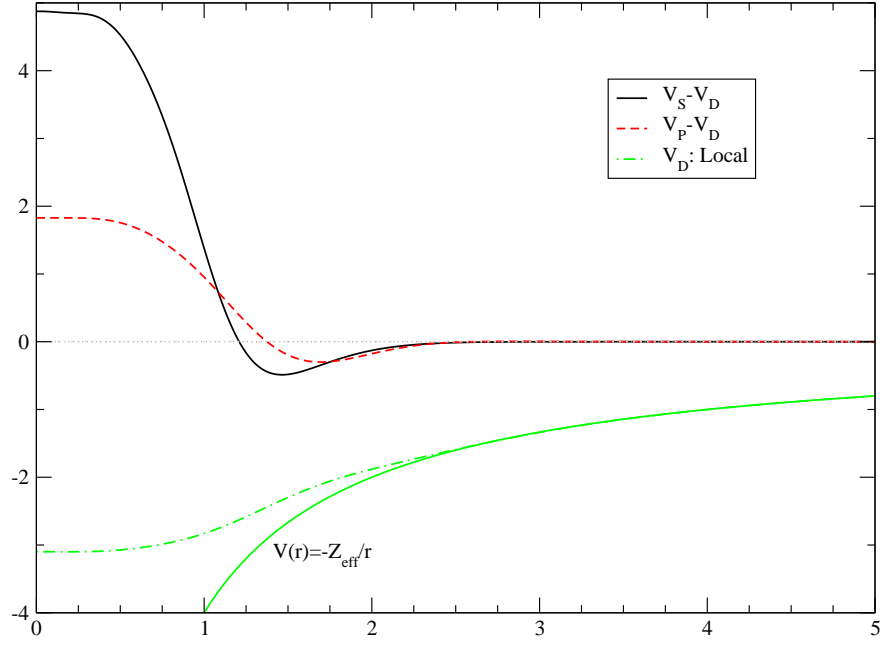


Figure 7.1: The Ge pseudopotential. The d channel is taken as the local component: the non-local corrections are $(V_s - V_d)$ and $(V_p - V_d)$. All units in a.u.

7.6 Pseudopotentials in Quantum Monte Carlo

7.6.1 Introduction

Current QMC methods are largely intractable for heavy atoms due to the poor scaling of the computation time with respect to the number of electrons. The core electrons contribute most to the total energy and require smaller time steps to properly sample, due to the small variations of the wave function near the nucleus. For an atom of atomic number Z , estimates for the scaling range from $Z^{5.5}$ [33] to $Z^{6.5}$ [7] for all-electron calculations. Recently Ma *et al* [36] estimated the scaling to be of the order $Z^{5.47}$.

7.6.2 Evaluating Non-Local Pseudopotentials in QMC

Excellent references for implementing Non-Local Pseudopotentials (NLPP)s in QMC can be found in Refs. [153, 210, 211, 31, 7]. For the sake of simplification the system under consideration is a single atomic core located at the origin. The results can easily be generalized to multiple atomic cores located at the positions $\{\mathbf{R}_\alpha\}$. The generic form of a non(semi)-local pseudopotential included in the valence only Hamiltonian for an atom, from section 7.2, and repeated here is

$$\hat{V}^{PP} = \sum_{lm} |Y_{lm}\rangle V_l(r) \langle Y_{lm}|. \quad (7.14)$$

It is conventional to divide the pseudopotential for an atom into a local part V_{loc} , common for all angular momentum, and a non-local correction to V_{loc} for each angular momentum channel $\hat{V}_{NL,l}$. The local part is typically chosen to be the largest angular momentum component, and includes the long range $-Z_{eff}/r$ interaction, where Z_{eff} is the effective atomic charge. The non-local corrections are relatively short ranged, with an extent of only a few bohr radii

around the atomic center. We now write the NLPP as

$$\begin{aligned}\hat{V}^{PP} &= \sum_i [V_{loc}(r_i) + |Y_{lm}\rangle(V_l(r_i) - V_{loc}(r_i))\langle Y_{lm}|] \\ &= \sum_i [V_{loc}(r_i) + \hat{V}_{NL,i}].\end{aligned}\quad (7.15)$$

The effect of the non-local operator $\hat{V}_{NL,i}$ on an arbitrary function of r_i is

$$\hat{V}_{NL,i}f(\mathbf{r}_i) = \sum_{l,m} v_l(r_i)Y_{lm}(\Theta_i) \int d\Omega'_i [Y_{lm}(\Omega'_i)]^* f(\mathbf{r}'_i), \quad (7.16)$$

where Ω_i is the angular part of the spherical coordinate \mathbf{r}_i and $v_l = V_l - V_{Loc}$. The effect of the angular momentum integral is to project the different angular momentum channels (s, p, d, \dots) out of $f(\mathbf{r}_i)$: each channel feels its own effective potential.

The contribution of \hat{V}_{NL} to the local energy $E_L = \Psi_T^{-1}H\Psi_T$ is

$$\Psi_T^{-1}\hat{V}_{NL}\Psi_T = \sum_i \Psi_T^{-1}\hat{V}_{NL,i}\Psi_T = \sum_i V_{NL,i}, \quad (7.17)$$

where the contribution from the i th electron is

$$V_{NL,i} = \sum_l v_l(r_i) \sum_{m=-l}^l Y_{lm}(\Omega_i) \int d\Omega'_i [Y_{lm}(\Omega'_i)]^* \frac{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}'_i, \dots, \mathbf{r}_N)}{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)}. \quad (7.18)$$

The angular integration is over a sphere centered at the atomic core and passing through the i th electron. Eq. (7.18) can be simplified by noting that $Y_{lm}(0,0) = 0$ for $m \neq 0$, and orienting the z -axis along \mathbf{r}_i to yield

$$V_{NL,i} = \sum_l v_l(r_i)Y_{l0}(0,0) \int d\Omega'_i [Y_{lm}(\Omega'_i)]^* \frac{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}'_i, \dots, \mathbf{r}_N)}{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)}, \quad (7.19)$$

here $\Omega_{r'}$ is the angular part of the vector \mathbf{r}' when \mathbf{r} points along the z -axis. It is convenient

to rewrite Eq. (7.19) in terms of the Legendre polynomials

$$P_l[\cos \gamma] = \frac{4\pi}{2l+1} \sum_{m=-l}^l [Y_{lm}(\theta', \phi')]^* Y_{lm}(\theta, \phi),$$

where γ is the angle between the vectors \mathbf{r} and \mathbf{r}' ,

$$V_{NL,i} = \sum_l \frac{2l+1}{4\pi} v_l(r_i) \int d\Omega'_i P_l[\cos(\theta'_i)] \frac{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}'_i, \dots, \mathbf{r}_N)}{\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)}. \quad (7.20)$$

The \mathbf{r}' dependence of Ψ_T has a primarily angular momentum character of the orbitals in the Slater determinant. The Jastrow correlation part of Ψ_T also has an angular momentum dependence and may introduce higher angular momentum components than the determinant part.

In principle, the expression for the energy due to electron i and the atom, Eq. (7.20), needs to be summed over all the electrons to give the correct total non-local energy in Eq. (7.15). However, since the non-local potentials are short ranged only the electrons within a sphere determined by the maximum extent of $v_l(r)$ need to be included in the summation.

The angular integral in Eq. (7.20) must be performed numerically. A quadrature rule on a sphere is chosen such that the integration of products of spherical harmonics is exact up to a maximum angular momentum l_{max} such that

$$\sum_{i=1}^N W_i Y_{l0}^*(\Omega_i) Y_{l'm}(\Omega_i) = \delta_{l'l} \delta_{0m}, \quad (7.21)$$

for $l \leq l_{max}$. In the above equation the N grid points on the sphere are $\{\Omega_i\}$ and their associated weights $\{W_i\}$.

It is also necessary to perform a random rotation of the grid every time the integral is evaluated. The coordinate axes are rotated such that the z -axis has a uniform solid angle distribution. This has the effect of ensuring that the estimator of any function evaluated on the sphere is unbiased [153]. The rotation of the coordinate system is defined by the Euler

Table 7.2: The spherical grid points and their associated weights: the grid points correspond to the vertices of a regular icosahedron (12 points) on a sphere of radius unity. This is the grid used to evaluate the non-local pseudopotentials in this thesis, including Ge and Si.

	x	y	z	weight
1	0.000000	0.525731	0.850651	0.083333
2	0.000000	-0.525731	-0.850651	0.083333
3	0.000000	-0.525731	0.850651	0.083333
4	0.000000	0.525731	-0.850651	0.083333
5	-0.850651	0.000000	0.525731	0.083333
6	0.850651	0.000000	-0.525731	0.083333
7	0.850651	0.000000	0.525731	0.083333
8	-0.850651	0.000000	-0.525731	0.083333
9	-0.525731	0.850651	0.000000	0.083333
10	0.525731	-0.850651	0.000000	0.083333
11	0.525731	0.850651	0.000000	0.083333
12	-0.525731	-0.850651	0.000000	0.083333

angles, which can be found in Ref. [212].

For the Ge pseudopotential the appropriate grid is constructed from the vertices of a regular icosahedron (12 points) with constant weights [153]. This grid, see Table 7.2, is exact for integration of spherical harmonics with $l \leq 5$.

7.6.3 The Consequences of Using Non-Local Pseudopotentials in Diffusion Monte Carlo

The inclusion of a non-local pseudopotential within DMC results in a sign problem analogous to the fermion sign problem (see Section 4.11.1). The origin of this problem can be traced to the imaginary time Schrodinger equation for a Hamiltonian containing a non-local pseudopotential, $H = T + V_{Loc} + V_{NL}$. The corresponding Greens function in the short time approximation has a term of the form $\langle \mathbf{R}' | \exp(-\tau V_{NL}) | \mathbf{R} \rangle$. The relative sign of this term depends on the values of \mathbf{R}' , \mathbf{R} and τ . Consequently, the Greens function can no longer be interpreted as a transition probability density.

The importance sampling transformation, Eq. (4.64), was made under the assumption

that the potential V is a local function. The transformation for the case where V includes a non-local part:

$$-\Psi_T(\mathbf{R}) \frac{\partial \Phi(\mathbf{R}, \tau)}{\partial t} = \Psi_T(\mathbf{R}) \left[-\frac{1}{2} \nabla^2 + V_{Loc} + V_{NL} - E_T \right] \Phi(\mathbf{R}, \tau) \quad (7.22)$$

can be shown to reduce to

$$\begin{aligned} -\frac{\partial f(\mathbf{R}, t)}{\partial t} = & -\frac{1}{2} \nabla^2 f(\mathbf{R}, t) + \nabla \cdot [\mathbf{V}_D(\mathbf{R}) f(\mathbf{R}, t)] + [E_L(\mathbf{R}) - E_T] f(\mathbf{R}, t) \\ & + \left\{ \frac{V_{NL} \Phi(\mathbf{R}, \tau)}{\Phi(\mathbf{R}, \tau)} - \frac{V_{NL} \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \right\} f(\mathbf{R}, \tau), \end{aligned} \quad (7.23)$$

where $f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R}) \Phi(\mathbf{R}, \tau)$. The last term in curly brackets includes the operation of a non-local potential on an unknown wave function and represents a non-local branching. In the *locality* approximation [210] this term is neglected such that Eq. (7.23) becomes identical to the original importance sampling transformation derived for local potentials

$$-\frac{\partial f(\mathbf{R}, t)}{\partial t} = -\frac{1}{2} \nabla^2 f(\mathbf{R}, t) + \nabla \cdot [\mathbf{V}_D(\mathbf{R}) f(\mathbf{R}, t)] + [E_L(\mathbf{R}) - E_T] f(\mathbf{R}, t). \quad (7.24)$$

The DMC solution to Eq. (7.24) is the stationary distribution

$$\lim_{\tau \rightarrow \infty} f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R}) \Phi_A(\mathbf{R}), \quad (7.25)$$

which is used to determine the mixed estimate of the energy E_A (see Section 4.10.4). The wave function Φ_A and the corresponding energy E_A are exact in the limit where the trial function is equal to the exact ground state eigenfunction Φ_0 , since as $\Psi_T \rightarrow \Phi_0$ the term in curly brackets approaches zero. An unfortunate consequence of the locality approximation is that the energy E_A is no longer variational; E_A is not an upper-bound to the exact energy E_0 , but is lower than the variational energy.

Table 7.3: Test results for VMC calculations of atoms and molecules featuring pseudopotentials. The trial wave function for VMC is identical to the HF wave function and does not include a Jastrow correlation factor. All HF results are from the Gaussian03 package [11]. The Ge pseudopotential from Ref. [12] is non-relativistic; the atomic basis is the uncontracted set $6s6p1d$. The C pseudopotential and basis are from Refs. [13, 14]. The methane molecule uses the same C pseudopotential and basis as the C atom. The H potential is the bare $-1/r$ and the H basis set $12s1p$ is contracted to $3s1p$. The C-H bond length is 2.06738 Bohr. The second column is the number of core electrons replaced by the pseudopotential (N_{Core}) and the third column is the total number valence electrons ($N_{Valence}$). The reason why the error bar for the Ge calculation is smaller than the calculations involving C is that the Ge pseudopotential has $l_{max} = 2$, while C has $l_{max} = 3$. All energies in a.u.

	N_{Core}	$N_{Valence}$	l_{max}	HF	VMC
C(3P)	2	4	3	-5.32233	-5.32199(41)
Ge(3P)	28	4	2	-3.628901	-3.628939(89)
CH ₄	2	8	3	-7.83536	-7.83526(73)

7.7 Core Polarization Potentials

The physics of the core polarization potential (CPP) can be traced to valence electrons inducing a core polarization and feeling the induced potential. The energy of a core-dipole induced by an electric field acting on the core is added to the valence Hamiltonian H_{val} . In a point dipole picture, the electric field acting on core J can be written in terms of the positions of the cores $\{\mathbf{R}_J\}$ and the valence electrons $\{\mathbf{r}_i\}$:

$$\mathbf{E}_J = \sum_i \frac{\mathbf{r}_{Ji}}{r_{Ji}^3} - \sum_{J' \neq J} \frac{\mathbf{R}_{JJ'}}{R_{JJ'}^3} Z_{J'} = \mathbf{E}_J^e + \mathbf{E}_J^I, \quad (7.26)$$

where $r_{Ji} = |\mathbf{r}_i - \mathbf{R}_J|$ is the electron-core distance and $R_{JJ'} = |\mathbf{R}_{J'} - \mathbf{R}_J|$ is the core-core distance. The full CPP is given by

$$\mathcal{H}_{CPP} = -\frac{1}{2} \sum_J \alpha_J \mathbf{E}_J \cdot \mathbf{E}_J, \quad (7.27)$$

where α_J is the static polarizability of core J . Inserting Eq. (7.26) into Eq. (7.27) yields

$$\begin{aligned} \mathcal{H}_{CPP} = & - \frac{1}{2} \sum_J \alpha_J \left\{ \sum_i \frac{1}{r_{Ji}^4} + \sum_{i \neq j} \frac{\mathbf{r}_{Ji} \cdot \mathbf{r}_{Jj}}{r_{Ji}^3 r_{Jj}^3} \right. \\ & \left. - 2 \sum_i \sum_{J' \neq J} \frac{\mathbf{r}_{Ji} \cdot \mathbf{R}_{JJ'}}{r_{Ji}^3 R_{JJ'}^3} Z_{J'} + \left| \sum_{J' \neq J} \frac{\mathbf{R}_{JJ'}}{R_{JJ'}^3} Z_{J'} \right|^2 \right\}, \end{aligned} \quad (7.28)$$

$$\mathcal{H}_{CPP} = -\frac{1}{2} \sum_J \alpha_J [V_e^J + V_{e-e}^J + V_{e-I}^J + V_{I-I}^J]. \quad (7.29)$$

The first two terms represent a single and two-electron interaction, respectively, and will be discussed in more detail. The third term reflects the indirect interaction of a single valence electron and another core. For a valence electron close to core J' , $\mathbf{r}_{Ji} \approx \mathbf{R}_{JJ'}$, therefore the third term is repulsive in character [40]. Finally, the last term is a pure core-core interaction term and is a constant.

The core dipoles are not point dipoles so appropriate corrections must be made when valence electrons penetrate the cores. The electric field acting on core J due to the valence electrons, \mathbf{E}_J^e in Eq. (7.26), is truncated at close proximity to the core center by a cutoff function. Shirley and Martin implemented the function introduced by Müller, Flesch and Meyer [40] in their study of alkali and alkaline earth atoms:

$$f(r) = \left[1 - e^{-(r/\bar{r})^2} \right]^2, \quad (7.30)$$

where \bar{r} is an adjustable parameter. This function approaches zero at the origin and has no effect at long range.

The form of the one-electron term used by SM has an explicit angular momentum dependence

$$\hat{V}_e = -\frac{1}{2} \alpha_J \sum_i \frac{1}{r_{Ji}^4} \sum_l f^2 \left(\frac{r_{Ji}}{\bar{r}_J^l} \right) \hat{P}^l, \quad (7.31)$$

where l is the angular momentum and \hat{P}^l is the projector of the l th angular momentum of

electron i with respect to the J th ion core. At long range this term has the form $-\alpha/2r^4$, as required by Born-Heisenberg result [213] for the interaction of an electron and a dipole of polarizability α . The parameters $\{\bar{r}_J^l\}$ are fit for the s , p and d channels by forcing \hat{V}_e^l and the self-energy, calculated in a generalized GW approximation, to have the same expectation values for the lowest valence states of each angular momentum in bare-core configurations. For one electron outside a core the valence properties are determined by the one-electron Green's function, which is a solution to Dyson's equation. All core-valence interactions beyond the HF level are incorporated into the self-energy: the self energy corrects Koopman's eigenvalues.

The two electron term has an analogous form

$$\hat{V}_{e-e} = -\frac{1}{2} \sum_J \alpha_J \sum_{i \neq j} \frac{\mathbf{r}_{Ji} \cdot \mathbf{r}_{Jj}}{r_{Ji}^3 r_{Jj}^3} f\left(\frac{r_{Ji}}{\bar{r}_J}\right) f\left(\frac{r_{Jj}}{\bar{r}_J}\right), \quad (7.32)$$

where \bar{r} is assigned the average of the s and p channels

$$\bar{r} = \frac{\bar{r}^s + \bar{r}^p}{2} \quad (7.33)$$

for $s - p$ bonding systems. This term screens the dipole exchange [43] and reduces the valence-valence correlation effects because correlated electrons tend to be on opposite sides of a given core due to their Coulomb repulsion while the factor $\mathbf{r}_{Ji} \cdot \mathbf{r}_{Jj}$ discourages such correlation.

7.7.1 Core Polarization Potential for Germanium

The evaluation of the SM CPP can be greatly simplified by making the \hat{V}_e term *local*. This is justified for Ge since the cutoff radii $\{\bar{r}^l\}$ for the s , p and d channels are approximately equal, see Table 1 in Ref. [43] for the CPP parameters for use in relativistic HF pseudopotentials. For Ge we use the core polarizability $\alpha = 0.7732$ and a single cutoff radius, $\bar{r} = 0.9063$, which

is the angular momentum weighted average of \bar{r}^0 and \bar{r}^1 . The CPP can now be evaluated by inserting the effective electric field

$$\bar{\mathbf{E}}_J = \sum_i \frac{\mathbf{r}_{Ji}}{r_{Ji}^3} f\left(\frac{r_{Ji}}{\bar{r}}\right) - \sum_{J' \neq J} \frac{\mathbf{R}_{JJ'}}{R_{JJ'}^3} Z_{J'} \quad (7.34)$$

into Eq. (7.27)². The inclusion of the CPP within QMC has little effect on the computation time since most of the required quantities are already necessary to evaluate the trial wave function and the electron-ion and electron-electron interaction energies. Figure 7.2 plots the \hat{V}_e term of the CPP.

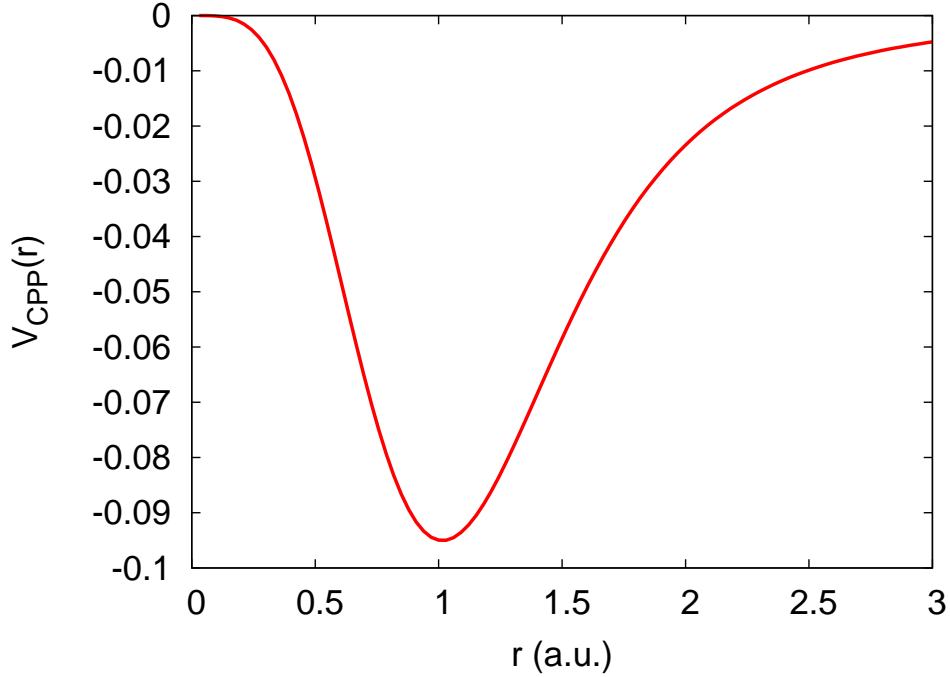


Figure 7.2: The V_e term of \mathcal{H}_{CPP} (see Eq. (7.31)) for the Ge atom.

7.7.2 Core Polarization Potentials in QMC: Previous Work

CPPs have successfully been included in QMC by several authors. The first known use of a CPP in QMC was by Shirley *et al.* [44] for calculations of the ionization potential and

²A more general relation for the case when \mathcal{H}_{CPP} is semi-local can be found in Ref. [45].

electron affinity of the Na atom and the bond length and disassociation energy of the Na_2 dimer. Maezono *et al.* [45] also studied Na, replacing the Na^+ core with a relativistic HF pseudopotential and modeled the core-valence correlations by the SM CPP. The inclusion of the CPP greatly improved the ionization energy of the atom, while there was marginal improvement for the cohesive energy of the solid. For the periodic system the electric field was evaluated as an analytic derivative of the Ewald potential. Alfè *et al.* [46] estimated that the inclusion of the CPP reduced the pressure induced phase transition from the diamond to β -tin structure of Si from 17.7 GPa to 16.5 GPa. Lee and Needs [47] developed their own parameterized CPPs for Si and Ti that use a different cutoff function $f(r)$ and are angular momentum independent, in contrast to the SM CPPs. This work was motivated by an earlier study of Ti that found the SM CPP is much too attractive in the d -channel, and leads to worse results in valence correlated wave function calculations than if the CPP was neglected [207]. The CPP parameters were fit such that the inclusion of the CPP to a relativistic HF pseudopotential reproduced the core-valence correlation energies calculated using a multiconfiguration HF (MCHF) method. Only double substitutions from the HF reference state are included, consisting of one core and one valence substitution. Systematic comparisons between the Lee and Needs CPP [47] and SM CPP for atomic excitations found that the SM CPP performed slightly better in Si, while in Ti the Lee and Needs CPP worked much better. Finally, in a study that is directly relevant to this work, Porter *et al.* [26] included the CPP in DMC calculations of the optical properties of the Si_nH_m nanoclusters, for $(n + m) \leq 34$.

Chapter 8

Results for Atomic Systems: Effect of the CPP on Atomic Excitations of Ge

8.1 Introduction

Quantifying the effect of core-valence correlation on atomic excitations is a natural starting point for examining the role of core-valence correlation on the excitations of larger Ge structures, including the Ge_nH_m molecules and nanoclusters and even bulk Ge. Core-valence correlation is incorporated by the SM CPP which for atoms includes the single-electron \hat{V}_e and electron-electron \hat{V}_{e-e} interactions in Eq. (7.29). These calculations are also useful in determining what role, if any, the relative occupation of the s and p atomic orbitals has on the size of the core-valence correlation energy.

Farid and Needs [214] have performed calculations for the first four ionization energies of Ge and calculated the total energy of the pseudoatom. The purpose of these calculations was to compare the quality of two norm-conserving relativistic LSDA pseudopotentials, one including non-linear core-corrections and the other one not, to all-electron and experimental results. The quality of the pseudopotential results was found to be much poorer for Ge than for Al, Si and C; for the other elements in the study the pseudopotentials including the non-linear core corrections were found to produce more accurate results. It should be noted that older experimental values, which can be found in Ref. [15], are used for comparison in this paper.

Shirley [17, 134] has performed HF and GW calculations on Ge; the parameterized CPP is fit to GW results for a single valence electron of each angular momentum channel outside a bare core.

8.2 Computational Method

The Quantum Monte Carlo (QMC) results are from the Variational Monte Carlo (VMC) method and the Diffusion Monte Carlo (DMC) method within the fixed-node approximation (FN-DMC) as discussed in Chapter 4. VMC explicitly includes correlation by using a many-body trial wave function Ψ_T . Fixed-node DMC is a stochastic method that uses the imaginary time propagator $e^{-\tau\mathcal{H}}$ to project the lowest energy state Φ_0 consistent with the nodal-surface of a trial wave function Ψ_T . The non-local pseudopotential is implemented within DMC by the locality approximation [210]. All QMC results are from the qmcPACK package [3].

The form of the trial wave function is the product of Slater determinants and Jastrow correlation factors:

$$\Psi_T(\mathbf{R}) = D^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) D^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N) \exp \left[\sum_{i<j} u(r_{ij}) + \sum_{iI} \chi_I(r_{iI}) \right]. \quad (8.1)$$

The Slater determinants for the spin-up electrons D^\uparrow and spin-down electrons D^\downarrow are constructed from single particle orbitals obtained from HF. The Jastrow electron-electron correlation factor u is spin-independent and has the form

$$u(r_{ij}) = ar_{ij}/(1 + br_{ij}), \quad (8.2)$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the electron-electron distance. The parameter $a = 0.5$ to satisfy the cusp condition for opposite spin electrons and b is adjustable. The cuspless electron-ion Jastrow factor has two adjustable parameters and is of the form

$$\chi_I(r_{iI}) = a_I/(1 + b_I r_{iI}^2), \quad (8.3)$$

where I is the ion index and $r_{iI} = |\mathbf{r}_i - \mathbf{R}_I|$ is the electron-ion distance. All of the Jastrow parameters are optimized within the VMC correlated sampling framework [155] by mini-

mizing a combination of the local energy and the variance of the local energy. Since the Jastrow terms are positive and multiplicative, the nodes of Ψ_T are determined by the Slater determinants and hence the accuracy of the FN-DMC results via the fixed-node error.

The Ge pseudopotential is discussed in Section 7.4, a CPP is added to this HF pseudopotential to account for core-valence correlations.

The Gaussian basis to describe the valence electrons of a Ge atom is the uncontracted set $5s5p1d$ [215]. The Gaussian exponents of this basis are optimized for each atomic configuration within HF. The quality of this basis is verified by the fact that the energy of the ground state $4s^24p^2(^3P)$ configuration is within a mHartree of the basis set limit [209], see Table 8.1. HF total energy calculations performed by the Gaussian03 [11] package for each of the atomic configurations provide the single particle orbitals for the Slater determinants

(
]

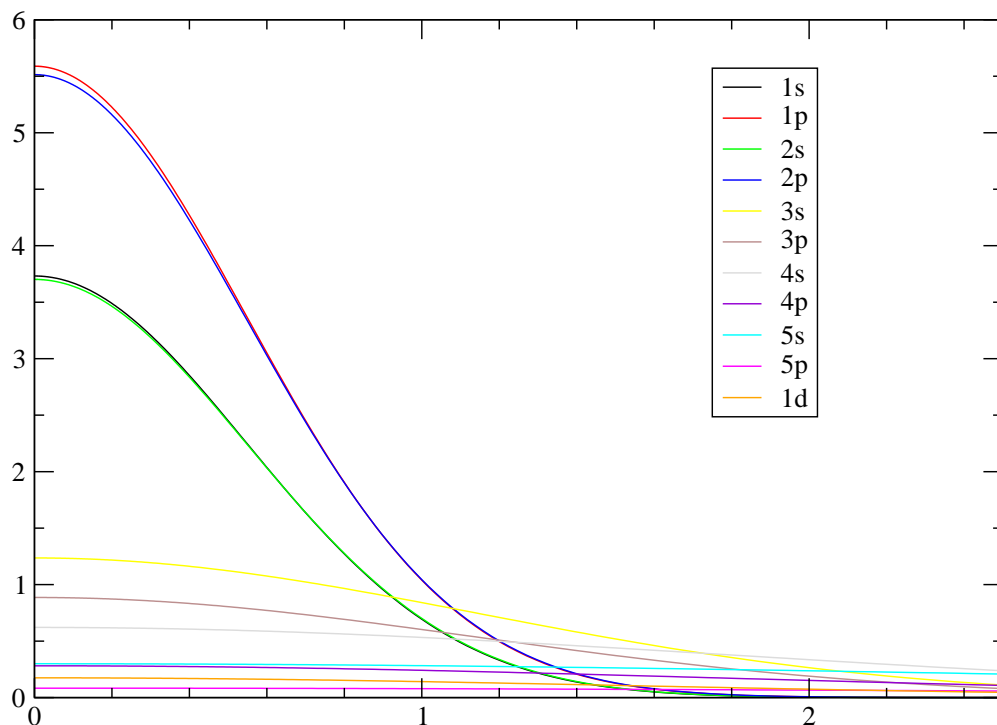


Figure 8.1: The basis for the ground state (3P) of the Ge atom.

Table 8.1: Comparison of the HF results for the $5s5p1d$ basis to the basis set limit (infinite basis) for the (3P) ground-state of the Ge atom. All energies in a.u.

Basis	$5s5p1d$	Infinite-Ref. [209]
Total Energy	-3.66625	-3.66673

Table 8.2: The CPP estimator $\langle \mathcal{H}_{CPP} \rangle$ of various atomic configurations of Ge. All energies in eV.

No. of electrons	Configuration	$\langle \mathcal{H}_{CPP} \rangle$	
		VMC	DMC
4	$4s^24p^2(^3P)$	-1.74153(21)	-1.76550(68)
	$4s4p^3(^5S)$	-1.35562(26)	-1.37333(49)
3	$4s^24p(^2P)$	-1.68520(30)	-1.73214(68)
	$4s4p^2(^4P)$	-1.34206(26)	-1.31730(49)
2	$4s^2(^1S)$	-1.73938(30)	-1.75481(87)
1	$4s(^2S)$	-1.09059(14)	-1.10514(41)

8.3 Results for Excitations

Table 8.2 shows the results for the CPP estimator $\langle \mathcal{H}_{CPP} \rangle$ from VMC and DMC for several atomic configurations. The CPP estimator serves as a measure of the core-valence correlation energy. In VMC $\langle \mathcal{H}_{CPP} \rangle$ is evaluated by averaging the quantity $\Psi_T^{-1} \mathcal{H}_{CPP} \Psi_T$ with respect to the distribution $|\Psi_T|^2$, where Ψ_T is the trial wave function. In DMC the averaging is performed with respect to the mixed distribution $\Psi_T \Phi_0$, where Φ_0 is the lowest energy state consistent with the nodes of Ψ_T . The magnitude of $\langle \mathcal{H}_{CPP} \rangle$ are not noticeably different in VMC and DMC. In general, $\langle \mathcal{H}_{CPP} \rangle$ is largest for the configurations which have two electrons in a $4s$ -orbital, indicating an angular momentum dependence. The $4s^24p^2(^3P)$ ground-state and the $4s4p^3(^5S)$ excited state configurations in Table 8.2 both contain four electrons, but the magnitude of $\langle \mathcal{H}_{CPP} \rangle$ is larger for $4s^24p^2(^3P)$ by more than 25%. Consequentially, it can be expected that the size of the CPP correction for atomic excitations depends on the initial and final occupation of the s -orbitals.

Table 8.3 gives the QMC results, with and without the CPP, for several excitations of the Ge atom; including the s removal energy, the $s \rightarrow p$ promotion energy and the first four ionization potentials. We have also included the total energy of the pseudoatom with

and without the CPP. All excitation energies are evaluated as total energy differences; see Table 8.4 for the QMC results for the total energies. The CPP correction is calculated as the difference of the VMC (DMC) and VMC+CPP (DMC+CPP) results, where VMC+CPP (DMC+CPP) indicates inclusion of the CPP. The CPP correction can also be evaluated as the difference of the CPP estimators for the ground and excited states: $\langle \mathcal{H}_{CPP} \rangle - \langle \mathcal{H}_{CPP} \rangle$. As will be shown in the next section, these two methods of calculating the CPP correction are equivalent in VMC and are in very close agreement in DMC. The inclusion of the CPP increases the excitation energies and significantly improves both the VMC and DMC results relative to experiment. The best values being the DMC+CPP ones. Interestingly, the VMC+CPP results are in closer agreement to experiment than the DMC results for most of the excitations. From this we can conclude that the gain in energy due to the inclusion of core-valence correlations via the CPP are larger than the gain in energy due to the improved treatment of valence-valence correlations by DMC. As expected from Table 8.2, the CPP correction is largest for the excitations that involve removing or promoting an electron in a s -orbital, such as the s removal energy and $s \rightarrow p$ promotion energy. In contrast, the CPP correction is almost negligible for the first two ionization energies, which correspond to removing electrons in p -orbitals.

Previous QMC studies of Ge [49, 50, 51] have reported results for the total energy of the ground-state of the Ge pseudoatom. These calculations feature a local [52] LDA pseudopotential and use single-particle orbitals from LDA to construct the Slater determinant part of the trial wave functions. The best VMC result is -103.22(1) eV and the best DMC result is -103.42(3) eV [51].

Table 8.3 also includes HF results. This is mainly because we used the HF orbitals to construct the Slater determinants and to demonstrate how significantly QMC improves upon the HF results by treating the valence-valence interactions at a many-body level. It should be remembered that HF completely omits correlation and therefore tends to underbind atoms and the valence levels. This means that both the total energy and the valence eigenvalues

are not sufficiently negative [205]. In addition, GW results from Shirley [17] are included where available.

Table 8.3: For the Ge atom results for the $s \rightarrow p$ promotion energy, the s removal energy (s^-), the first four ionization energies and the total energy of the pseudoatom, with and without the CPP. All excitation energies are total energy differences. The CPP correction, ΔCPP , is calculated as the energy difference of the results with and without the CPP. Experimental results from Ref. [15] have been properly spin-orbit corrected [16]: the J weighted averaging has been performed with respect to the $4s^2(^1S)$ configuration, see Appendix A.10. All energies in eV and the QMC error bars are smaller 0.01 eV. GW results from Ref. [17].

	$s \rightarrow p$ $s^2p^2(^3P) \rightarrow$ $sp^3(^5S)$	s^- $s^2p^2(^3P) \rightarrow$ $sp^2(^4P)$
Expt.	5.08	14.31
DMC+CPP	5.12	14.37
DMC	4.74	13.94
ΔCPP	0.38	0.43
VMC+CPP	4.90	14.24
VMC	4.51	13.84
ΔCPP	0.38	0.40
GW	4.3	13.5
HF	3.48	12.28

	1st IP $s^2p^2(^3P) \rightarrow$ $s^2p(^2P)$	2nd IP $s^2p(^2P) \rightarrow$ $s^2(^1S)$	3rd IP $s^2(^1S) \rightarrow$ $s(^2S)$	4th IP $s(^2S) \rightarrow$ $[\text{Ar}]3d^{10}(^1S)$	Total Energy of $s^2p^2(^3P)$
Expt.	7.93	15.79	34.06	45.72	103.50
DMC+CPP	7.86	15.68	34.06	46.00	103.60
DMC	7.85	15.66	33.46	44.90	101.86
ΔCPP	0.02	0.01	0.60	1.11	1.74
VMC+CPP	7.89	15.43	34.02	45.98	103.31
VMC	7.83	15.48	33.38	44.89	101.57
ΔCPP	0.05	-0.05	0.65	1.09	1.74
HF	7.41	15.27	32.29	44.89	99.85

	HF	VMC	VMC+CPP	DMC	DMC+CPP
$(4s)^2(4p)^2$	-99.85330	-101.57461(54)	-103.31447(54)	-101.8640(15)	-103.6005(16)
$(4s)(4p)^3$	-96.37805	-97.06045(49)	-98.41686(49)	-97.11961(65)	-98.48026(73)
$(4s)^2(4p)$	-92.44232	-93.73979(79)	-95.42760(84)	-94.0159(16)	-95.7371(17)
$(4s)(4p)^2$	-87.57443	-87.73108(71)	-89.07211(68)	-87.9200(11)	-89.2270(12)
$(4s)^2$	-77.17685	-78.26320(57)	-80.00258(54)	-78.35362(95)	-80.06203(90)
$(4s)$	-44.88760	-44.88768(33)	-45.97826(35)	-44.89895(38)	-46.00457(41)

Table 8.4: Hartree-Fock (HF) and QMC total energies for various atomic configurations of Ge. All energies in eV.

Chapter 9

Results for Molecular Systems

9.1 Introduction

The importance of core-valence correlation for the atomic excitations of Ge provides a motivation to study the effect of the core polarization (CPP) on larger Ge structures. The next step in this thesis is to examine and quantify the effect of the CPP on the excitations of the closed-shell molecules GeH_4 and Ge_2H_6 . These molecules represent the molecular limit of the hydrogen passivated structures which have been the subject of several theoretical studies [58, 59, 60, 61, 62, 63, 64, 65, 1, 28]. The results for GeH_4 and Ge_2H_6 provide a benchmark for predicting the effect of the CPP on the larger structures Ge_5H_{12} , $\text{Ge}_{10}\text{H}_{16}$, $\text{Ge}_{14}\text{H}_{20}$ and $\text{Ge}_{29}\text{H}_{36}$, which are discussed in the following chapter. One purpose of the present calculations are to demonstrate that the CPP correction can be calculated from the CPP estimators instead of energy differences; this is useful for the larger clusters where it leads to a considerable saving in computer time. In addition, the results for excitation spectra can be compared with experiment and with corresponding results for SiH_4 and Si_2H_6 , which is useful for the analysis of the larger clusters given in the following chapter.

9.2 Computational Method

The Gaussian basis for the Ge atoms consist of the uncontracted set $4s4p1d$. This basis is nearly identical to the basis used for the atomic calculations (Section 8.2), with the exception that we removed a single diffuse sp shell. The H atomic basis is the primitive set $21s1p$ which

is contracted to $3s1p$. The 10 Gaussians which are contracted to form the $1s$ and $2s$ orbitals of H are fit to Slater type orbitals. This serves to better satisfy the cusp condition at the positions of the H nuclei, and has been found to significantly reduce the variance of the local energy $E_L = \Psi_T^{-1} \mathcal{H} \Psi_T$. For the Jastrow part of the trial wave function we used electron-electron and electron-nuclear correlations. The Jastrow part of Ψ_T contains a total of five adjustable parameters. All DMC calculations use a time-step of 0.05 a.u.

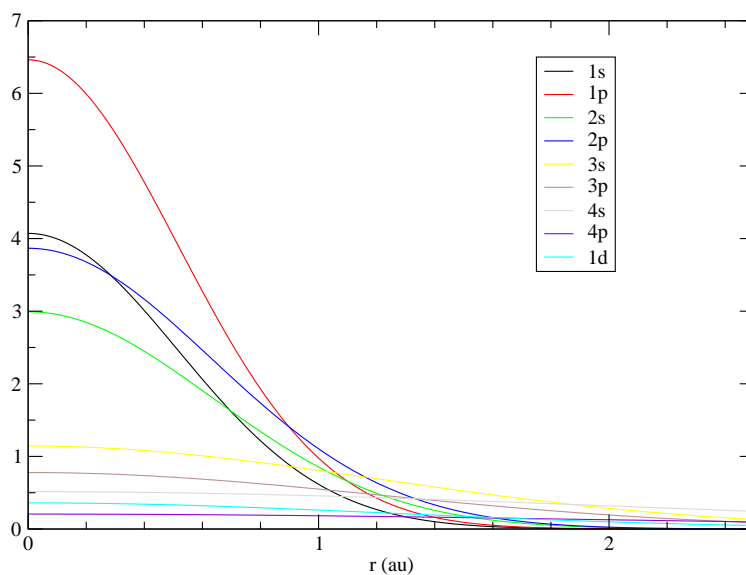
The discussion of the Ge pseudopotential and CPP found in Chapter 7. The H nuclei are represented by a $-1/r$ bare-core potential.

The optical excitations are calculated in QMC following the method in Section 4.13. Only the ground state trial wave function is optimized; optimization of the excited state wave function was found to have an insignificant effect on the optical gap.

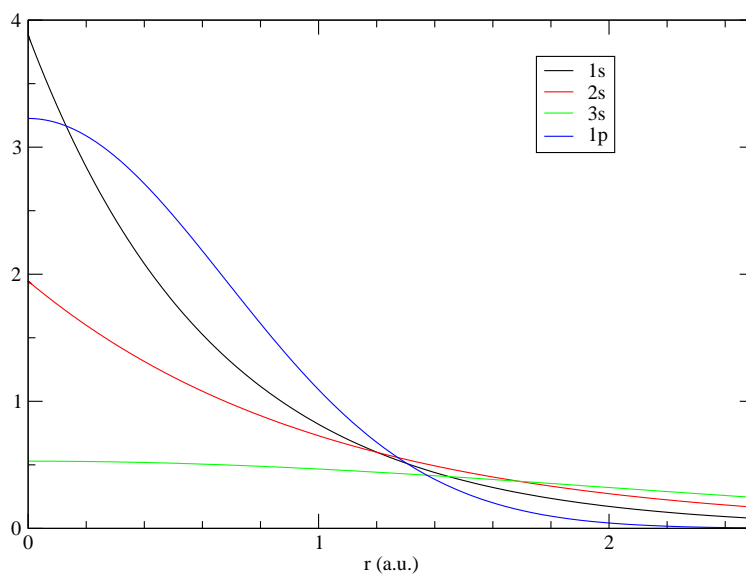
For the analysis of the effect of the CPP on the properties of GeH_4 and Ge_2H_6 , we chose to use the best available experimental data for the geometries. For the GeH_4 molecule, we use the bond length 1.520 Å from Ref. [216] and references therein. For Ge_2H_6 , we use the Ge-Ge bond length 2.403 Å, the Ge-H bond length 1.541 Å and Ge-Ge-H bond angle 112.5° [217], see Figure 9.2.

9.3 Results for Atomization Energies

We have obtained the atomization energies of 12.725(5) eV for GeH_4 and 20.812(6) eV for Ge_2H_6 in DMC with the addition of the CPP. The agreement with both experiment and other theoretical values is reasonable. Our DMC+CPP result for GeH_4 is close to the experimental value of 12.490(37) eV; this value has been extrapolated to 0 K and corrected for the zero-point energy and spin-orbit averaged for the Ge atoms [218]. For both GeH_4 and Ge_2H_6 we can compare directly with the atomization energies calculated previously by Ricca and Bauschlicher [219] using coupled-cluster singles and doubles with perturbative triples [CCSD(T)] methods which were extrapolated to attempt to reach the basis set limit.



(a) The atomic Ge basis.



(b) The atomic H basis.

Figure 9.1: The basis sets for the Ge atoms 9.1(a) and the H atoms 9.1(b). For H the linear combination of 10 Gaussians used to construct the $1s$ and $2s$ orbitals are each fit to a single STO.

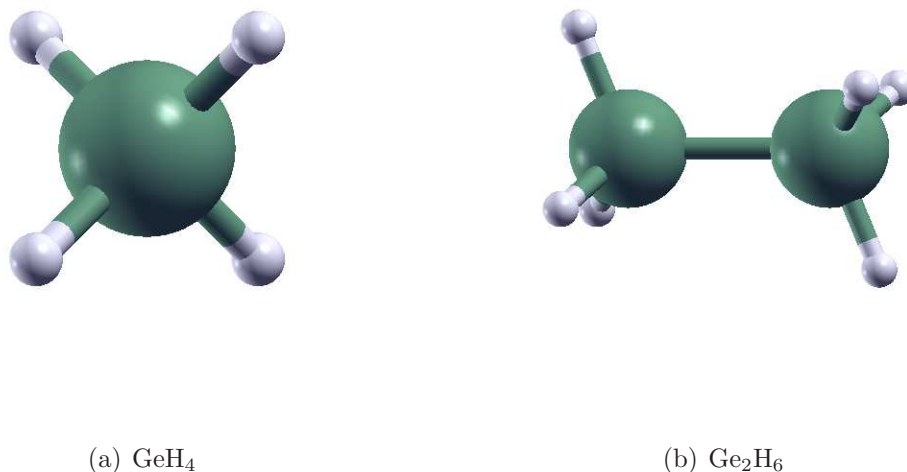


Figure 9.2: The GeH₄ and Ge₂H₆ molecules.

Using the same relaxed geometries found by those authors, our DMC+CPP atomization energies of 12.719(5) eV for GeH₄ and 20.800(6) eV for Ge₂H₆ are very close to the values of 12.78 eV and 20.99 eV found by Ricca and Bauschlicher.¹ Earlier work by Nicklass and Stoll [195] differs by ≈ 0.8 eV, which may be attributed to basis set convergence issues and differences in the geometries. Nonetheless, our work and that of Nicklass and Stoll conclude that core-valence correlation has only a small effect for the atomization energies since the magnitude of the core-valence interaction energy is roughly the same size for the molecule and the constituent atoms.

9.4 Treatment of the CPP: Estimating the CPP Correction to Excitations

One purpose of the calculations on the molecules is to show that it is possible can extract the CPP correction, ΔCPP , for excited state properties without having to perform two sets of

¹All comparisons are to the results from Table 5, excluding the spin-orbit and zero-point energy terms.

calculations, *i.e.* with and without the presence of the CPP in the valence Hamiltonian \mathcal{H}_{val} . This is especially important for the larger Ge_nH_m nanoclusters which involve considerably more computation time than the GeH_4 and Ge_2H_6 molecules to attain the same level of accuracy in the optical gap. Ideally, the CPP correction for excitations can be calculated as the difference of the CPP estimators of the ground and excited state

$$\Delta\text{CPP} \approx \Delta\langle H_{CPP} \rangle = |\langle \mathcal{H}_{CPP} \rangle^{EX} - \langle \mathcal{H}_{CPP} \rangle^{GS}|. \quad (9.1)$$

Hence, it is necessary to determine the accuracy of the CPP estimator in both the VMC and DMC methods. In VMC, the CPP estimator is equal to the difference of the energies obtained with and without the CPP

$$\langle \mathcal{H}_{CPP} \rangle = \langle \mathcal{H}_{val+CPP} \rangle - \langle \mathcal{H}_{val} \rangle, \quad (9.2)$$

since all of the averages $\langle \rangle$ are computed with respect to the probability distribution Ψ_T^2 . However, in the DMC method the estimators are computed with respect to the mixed distribution $f = \Psi_T\Phi_0$. The inclusion of the CPP in the valence Hamiltonian has the effect of changing the DMC solution: $\Phi_0 \mapsto \Phi_0^{CPP}$. Therefore, Eq. (9.2) is only an approximate relation in DMC, $\langle \mathcal{H}_{val+CPP} \rangle - \langle \mathcal{H}_{CPP} \rangle \approx \langle \mathcal{H}_{val} \rangle$. The terms on the left hand side of the equation are averaged with respect to the distribution $\Psi_T\Phi_0^{CPP}$, while the term on the right hand side is averaged with respect to $\Psi_T\Phi_0$.

Table 9.1 presents the comparison between the CPP estimator $\langle H_{CPP} \rangle$, and the energy difference $\langle \Delta E \rangle = \langle \mathcal{H}_{val+CPP} \rangle - \langle \mathcal{H}_{val} \rangle$, for the the total energy of the ground state and an excited state of the Ge_2H_6 molecule. (The excited state is created by promoting an electron from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO) in the spin-up Slater determinant). The VMC results are within an error bar of each other, while the DMC results are close enough to justify replacing ΔE with

$\langle H_{CPP} \rangle$. Table 9.1 also presents results for the accuracy of the CPP estimator method for the lowest singlet transition, which is the difference between the excited state energy and the ground state energy. For this case the comparison is between the difference of the CPP estimators of the excited state and ground state, see Eq. (9.1), and the difference of the energy of the singlet transition with and without the CPP in the valence Hamiltonian. Once again, the VMC results are within an error bar of each other, while the DMC results are within approximately 0.05 eV. Therefore, it can be concluded that it is possible to determine the CPP correction to excited state properties by subtracting the CPP estimators for the ground and excited states. All subsequent mentions to the CPP correction will be calculated by Eq. (9.1). Additionally, from Table 9.1 it can be seen that the magnitude of the CPP is larger for the excited states than for the ground state, indicating that the CPP should lower the energy of the optical excitations of Ge_2H_6 .

Table 9.1: The effect of the CPP (eV) on the ground and excited state energies of Ge_2H_6 calculated by two methods: (i) the differences in the QMC energies with and without the CPP $\langle \Delta E \rangle = \langle \mathcal{H}_{val+CPP} \rangle - \langle \mathcal{H}_{val} \rangle$ and (ii) the CPP estimator $\langle H_{CPP} \rangle$. The excited state is constructed by promoting a spin-up electron from the HOMO and placing it in the LUMO. Results are also presented for the effect of the CPP on the energy difference between the excited state and ground state, which is the singlet transition.

	CPP	Ground State (GS)	Excited State (EX)	EX - GS
VMC	$\langle \mathcal{H}_{CPP} \rangle$	-3.1976(21)	-3.4191(22)	-0.2215(30)
	$\langle \Delta E \rangle$	-3.203(17)	-3.412(18)	-0.209(25)
DMC	$\langle \mathcal{H}_{CPP} \rangle$	-3.3654(14)	-3.5849(14)	-0.2195(20)
	$\langle \Delta E \rangle$	-3.471(10)	-3.638(22)	-0.167(24)

9.5 Results for Excitations

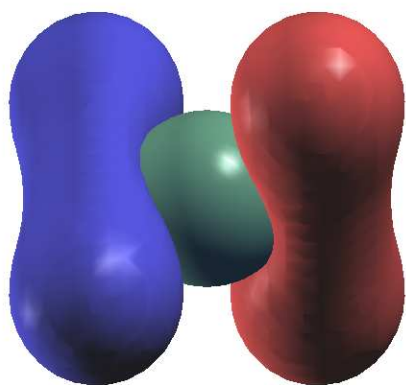
The main interest is determining the effect of core-valence correlation on the singlet transitions of GeH_4 and Ge_2H_6 , where the optical gap is the lowest dipole allowed singlet transition. Table 9.2 gives the VMC+CPP and DMC+CPP transition energies, as well as the change of energy due to the inclusion of the CPP, for the first two singlet transitions of GeH_4 , and the

first three singlet transitions of Ge_2H_6 . The Ge_2H_6 molecule provides a better example for studying the effect of core-valence correlation since it is the first of the Ge_nH_m structures to include a Ge-Ge bond. Each of these optical transitions involves removing an electron from the HOMO - t_2 for GeH_4 and A_{1g} for Ge_2H_6 - and promoting it to an excited state molecular orbital, which are ordered by their single particle energies. The QMC gaps are considerably lower than the HF gaps, with the DMC+CPP gaps being the lowest. The inclusion of the CPP has the effect of *lowering* the transition energies.

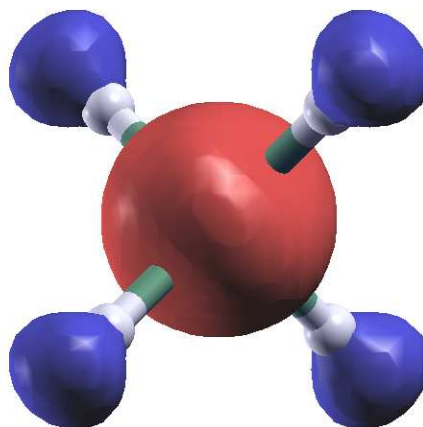
Table 9.2: The QMC results (eV) for the first two singlet transitions of GeH_4 and the first three singlet transitions of Ge_2H_6 , with the corresponding CPP corrections ΔCPP . The transition energies are calculated by Eq. (4.96), where the excited state is created by promoting an electron from the HOMO to the excited state orbital denoted by its irreducible representation. The transition corresponding to the optical gap is in bold. We have used the experimental geometries for GeH_4 and Ge_2H_6 .

	GeH ₄ [T_d] HOMO t_2		Ge ₂ H ₆ [D_{3d}] HOMO A_{1g}		
Excited State	a_1	t_2	A_{2u}	E_u	A_{1g}
HF	17.719	18.372	15.006	15.155	15.786
VMC+CPP	10.266(12)	10.576(12)	7.976(18)	8.088(17)	9.146(19)
ΔCPP	-0.2488(20)	-0.0944(20)	-0.2215(30)	-0.0185(44)	-0.2242(33)
DMC+CPP	9.479(13)	10.110(11)	7.541(12)	7.922(11)	8.169(15)
ΔCPP	-0.2221(18)	-0.0732(17)	-0.2195(20)	-0.0154(19)	-0.1994(19)

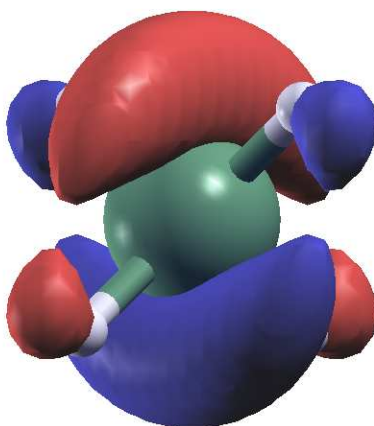
The size of the CPP correction strongly depends on the symmetry of the excited state. For GeH_4 , the size of the CPP correction is much larger for the $t_2 \rightarrow a_1$ transition than the optically forbidden $t_2 \rightarrow t_2$ transition. This is also the case for Ge_2H_6 : the CPP correction is largest for the $A_{1g} \rightarrow A_{2u}$ and $A_{1g} \rightarrow A_{1g}$ transitions, while it is almost negligible for the $A_{1g} \rightarrow E_u$ transition. The difference in the CPP corrections can be explained by the relative s to p Ge atomic orbital character of the excited state molecular orbitals. For GeH_4 , the a_1 molecular orbital has a s character of the Ge atomic orbitals while the t_2 has a p character. For Ge_2H_6 , the A_{1g} and A_{2u} molecular orbitals have a s character of the Ge atomic orbitals, while the E_u molecular orbital has a character of the Ge p atomic orbitals. As was demonstrated for the Ge atom in Chapter 8, the size of the core-valence correlation energy is



(a) HOMO - t_2

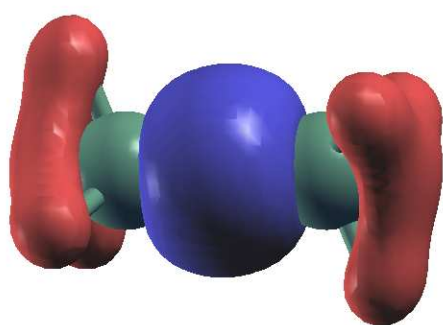


(b) LUMO - a_1

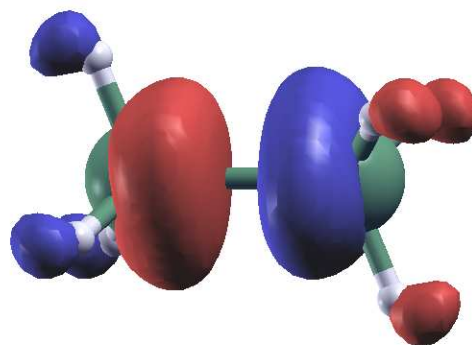


(c) 2nd LUMO - t_2

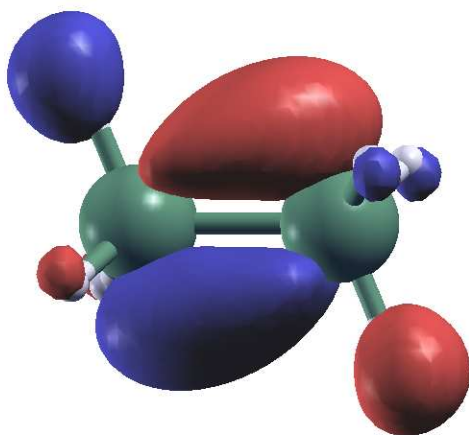
Figure 9.3: Isosurfaces for several molecular orbitals of GeH_4 , including the HOMO and the first two excited orbitals (red +, blue -). The t_2 MO is triply degenerate. The irreducible representation is also denoted.



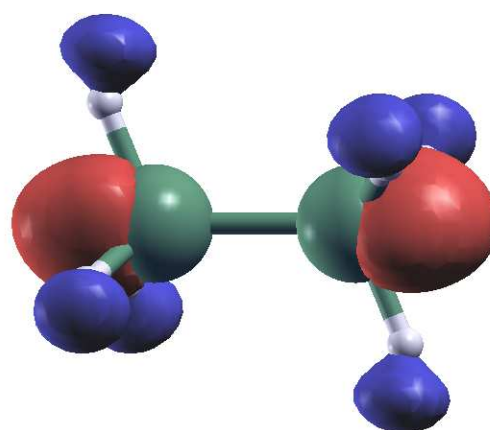
(a) HOMO - A_{1g}



(b) LUMO - A_{2u}



(c) 2nd LUMO - E_u



(d) 3rd LUMO - A_{1g}

Figure 9.4: Isosurfaces for several molecular orbitals of Ge_2H_6 , including the HOMO and the first three excited orbitals (red +, blue -). The E_u MO is doubly degenerate. The irreducible representation is also denoted.

different for different angular momentum channels, and is a larger for electrons in s -orbitals than in p -orbitals. Figures 9.3 and 9.4 plot the isosurfaces of the molecular orbitals used to evaluate the transition energies in Table 9.2 for GeH_4 and Ge_2H_6 :² from the isosurfaces it is relatively easy to visualize the relative Ge atomic s and p character of the molecular orbital. Table 9.3 provides the total energies of the ground and excited states of GeH_4 and Ge_2H_6 used in constructing Table 9.2, while Table 9.4 provides the CPP estimators.

Table 9.3: Hartree-Fock (HF) and QMC total energies for the ground-state E^{GS} and excited states E^{EX} of GeH_4 and Ge_2H_6 : the subscript on E^{EX} denotes the unoccupied molecular orbital to which the spin-up electron is promoted. The best available experimental geometries are used. All energies in eV.

(a) GeH_4					
	HF	VMC	VMC+CPP	DMC	DMC+CPP
E^{GS}	-163.8638	-167.9656(73)	-169.5953(73)	-169.0029(71)	-170.7480(82)
E_{LUMO}^{EX}	-146.1445	-157.4558(93)	-159.3293(90)	-159.369(12)	-161.269(11)
$E_{2ndLUMO}^{EX}$	-145.4920		-159.0196(93)		-160.6384(82)

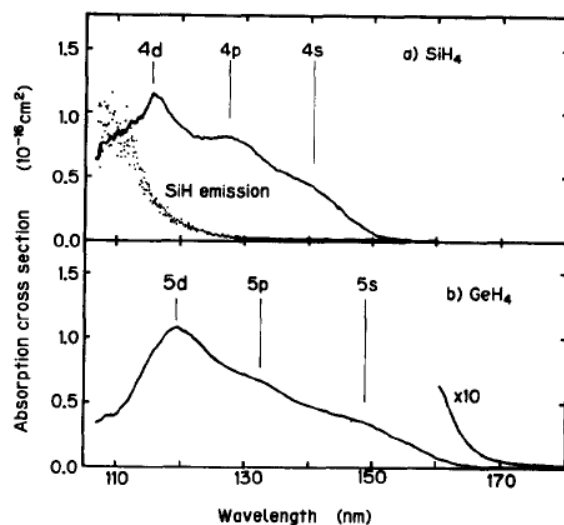
(b) Ge_2H_6					
	HF	VMC	VMC+CPP	DMC	DMC+CPP
E^{GS}	-296.7392	-304.485(11)	-307.688(12)	-306.1756(68)	-309.6470(76)
E_{LUMO}^{EX}	-281.7332	-296.300(11)	-299.712(13)	-298.468(20)	-302.1064(90)
$E_{2ndLUMO}^{EX}$	-281.5838		-299.600(12)		-301.7249(76)
$E_{3rdLUMO}^{EX}$	-280.9534		-298.542(14)		-301.481(13)

9.6 Comparison to Experiment

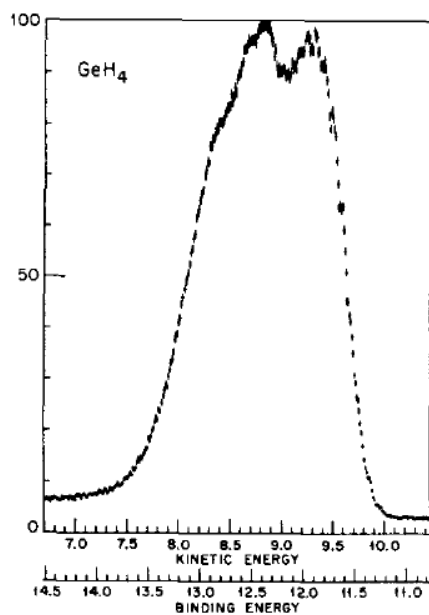
The results for GeH_4 can be compared with experiment and with previous QMC calculations for SiH_4 [48, 26, 57]. Because there is controversy and different interpretations of the experimental results, we will use the previous theoretical work on SiH_4 to make direct comparisons of the results and to aid in the comparison to experiment.

The absorption spectrum of GeH_4 has been reported to display broad peaks at 8.3 eV, 9.4 eV and 10.4 eV [19, 221, 222], although there is some disagreement about the location of

²All isosurfaces made using the XCrySDen package [220].



(a) The experimental absorption cross section for SiH_4 and GeH_4 from Itoh *et al.* [19]. The proposed Rydberg transitions are also indicated. The peaks for SiH_4 are located at 8.8 eV, 9.7 eV and 10.7 eV; GeH_4 the peaks are at 8.3 eV, 9.4 eV and 10.4 eV.



(b) The photoelectron spectrum of GeH_4 from Pullen *et al.* [18]. The peaks are located at 11.98 and 12.46 eV. The vertical scale is the counts/second.

Table 9.4: The CPP estimator $\langle \mathcal{H}_{CPP} \rangle$ for the ground-state and excited states of GeH_4 and Ge_2H_6 . The best available experimental geometries are used. The excited state is constructed by promoting a spin-up electron from the HOMO and placing it in the MO in parenthesis. All energies in eV.

(a) GeH_4		
	VMC	DMC
E^{GS}	-1.6273(12)	-1.6985(14)
E_{LUMO}^{EX}	-1.8761(16)	-1.9206(11)
$E_{2ndLUMO}^{EX}$	-1.7217(16)	-1.7718(9)

(b) Ge_2H_6		
	VMC	DMC
E^{GS}	-3.1976(21)	-3.3654(14)
E_{LUMO}^{EX}	-3.4191(22)	-3.5849(14)
$E_{2ndLUMO}^{EX}$	-3.2161(24)	-3.3808(13)
$E_{3rdLUMO}^{EX}$	-3.4219(26)	-3.5648(13)

the second peak [221]. The first peak has been assigned as a $t_2 \rightarrow a_1(5s)$ Rydberg transition, but there is disagreement in the assignment of the other peaks [19, 221]. The ionization energy is determined by the photoelectron spectrum, which displays two distinguishable peaks at 11.98 eV and 12.46 eV [18]. The same features are found in SiH_4 . The photoelectron spectrum of SiH_4 displays the same double peak structure as GeH_4 , with peaks at 12.36 eV and 12.85 eV [18]. The absorption spectrum of SiH_4 also displays a nearly identical structure to that of GeH_4 , with the peaks shifted to higher energies at 8.8 eV, 9.7 eV and 10.7 eV. Thus the overall conclusion is that both the ionization potential and the gaps are 0.3-0.5 eV higher in SiH_4 than in GeH_4 .

The DMC+CPP results for the ionization energy and the optical gap of GeH_4 are shown in Table 9.5 for both the experimental bond length (the same is in Table 9.2) and the calculated LDA bond length of 1.571 Å. The previous results for SiH_4 that are most comparable to our work are those of Porter *et al.* [48] who carried out a DMC+CPP calculation using the LDA bond length for SiH_4 . Their DMC+CPP result of 9.47(2) eV for the optical gap can be compared directly with the present result of 8.93(1) eV, *i.e.* a decrease in the gap

of 0.5 eV in going from SiH₄ to GeH₄. This is consistent with the difference in experiment between SiH₄ and GeH₄.

A full analysis of the optical structure and the experimental results requires an interpretation of the multiple peaks seen in experiment. Porter *et al.* [48] maintained, supporting earlier all-electron multi-reference single- and double-excitation configuration interaction (MRD-CI) work by Chantranupong *et al.* [223], that the first two peaks of the SiH₄ spectrum actually derive from a Jahn-Teller splitting of the triply degenerate $t_2 \rightarrow 4s$ transition, similar to the splitting in the photoelectron spectrum for the ionization energy [18].³ (The excited state involves partially filled degenerate t_2 HOMO levels, which is subject to Jahn-Teller distortion that can break the T_d symmetry and lift the degeneracy.) Thus the single calculated gap should be compared to the two experimental peaks, *i.e.* lie in between the two peaks, which is consistent with the calculated gaps for SiH₄ and GeH₄. This interpretation is different from Grossman *et al.* [57], whose result of 9.1(1) eV is closer to the first peak in the absorption spectrum. It is important to note that spin-orbit effects, which we have neglected, are expected to be considerably larger for GeH₄ than for SiH₄.

Table 9.5: DMC+CPP results (eV) compared to experiment for the ionization potential (IP) and optical gap (E_{gap}^{opt}) of GeH₄ in its experimental and LDA relaxed geometry. The IP is calculated as the total energy difference of the N and $N - 1$ electron systems: $IP = E[N] - E[N - 1]$. The optical gap is calculated by Eq. 4.96. The experimental data for the IP are from Ref. [18] and the optical gap from [19]. See the text for a discussion of the comparison of our single energy with the two peaks observed in experiment.

		Ge-H bond length:		
		Expt.	LDA	Expt.
IP	DMC+CPP	12.577(12)	12.4963(69)	11.98,12.46
	Δ CPP	-0.0456(19)	-0.0820(17)	
E_{gap}^{opt}	DMC+CPP	9.479(13)	8.934(13)	8.3,9.4
	Δ CPP	-0.2221(18)	-0.2272(21)	

³Several theoretical investigations have attempted to find the most stable structure of the GeH₄+ ionized molecule, although there is a disagreement about which is the correct structure [224, 225]. This is especially useful in determining the adiabatic ionization energy.

Chapter 10

Results for Nanoclusters

10.1 Introduction

In this chapter the optical gaps of the molecules and nanoclusters GeH_4 , Ge_2H_6 , Ge_5H_{12} , $\text{Ge}_{10}\text{H}_{16}$, $\text{Ge}_{14}\text{H}_{20}$ and $\text{Ge}_{29}\text{H}_{36}$ are presented (see Figure 10.11). The CPP approach outlined in Chapter 7 is used to treat the core-valence interactions. The computation details regarding the trial wave function, including the Ge and H basis sets, are identical to the description in Section 9.210.5.

10.2 The Structure of the Nanoclusters

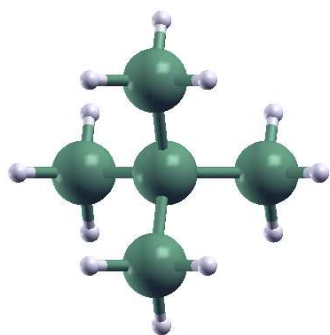
Nanostructured Ge and Si systems have been fabricated by a variety of techniques. Very small, nearly spherical Ge nanocrystals have been synthesized by rf cosputtering of Ge and SiO_2 [22], ion implantations [226], as colloidal nanocrystals [53] and reduction of metastable compounds [82]. It is typically assumed that these nanocrystals or nanoclusters retain a tetrahedral coordination in the center and that the surface dangling bonds are saturated, usually by hydrogen: ¹ these nanoclusters are known as the hydrogen terminated Ge or Si nanoclusters (Ge_nH_m or Si_nH_m). Unfortunately the exact ground state structures of all but the smallest Ge and Si nanoclusters are unknown. At the present time the biggest obstacle to commercial adaption of Ge or Si nanoclusters is the reliability of the fabrication process; it is difficult to prepare nanoclusters in a size and shape consistent manner, and without the

¹Hydrogen is present in many of the preparation procedures [53, 72]

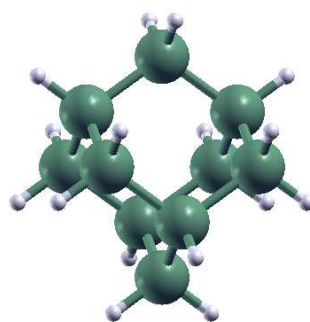
presence of contaminants.

The ground-state structures of all but the smallest Ge nanoclusters are unknown. Almost all the theoretical studies of Ge_nH_m and Si_nH_m have used LDA relaxed geometries derived from the bulk diamond structures of Ge and Si, as shown in Figure 10.1(f) for the $\text{Ge}_{29}\text{H}_{36}$ nanocluster. The motivation for this approach is twofold: (i) for large nanoclusters it is reasonable to expect that the core maintains bulk-like properties: (ii) bulk derived structures are ideal for studies that attempt to approach the bulk limit. Hydrogen is placed along the appropriate tetrahedral direction to saturate all the dangling bonds. The initial geometries of our clusters are determined by experimental values for the Ge-Ge bond length (4.65 Å) from the bulk and the Ge-H bond length (1.520 Å) from the GeH_4 molecule. Then, the structures are fully relaxed in DFT-LDA using an identical procedure to TM [1]. The LDA relaxation reduced the Ge-Ge bond by less than 1% and increased the Ge-H bond by 3-4%. We have not attempted to relax the geometries of the nanoclusters in their optically excited states to take the Stokes shift into account for the optical excitations.

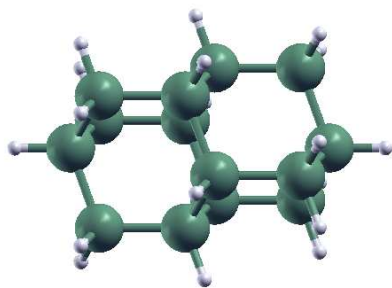
The fashion in which the nanoclusters are synthesized has been found to significantly effect the final structure that forms, and hence their optical properties. Draeger *et al.* [169] employed first-principles molecular dynamics (FPMD) and QMC to determine impact of the experimental synthesis process on structural and optical properties. This study was motivated by experiments that found optical gaps of Si clusters formed by physical vapor deposition several eV lower [227] than those formed using inverse micelles [228]. Draeger *et al.* found that the number of atoms in the core is larger than ideal bulk-derived structures when a finite temperature synthesis process is emulated. It was observed during the formation process of 1 nm clusters (approximately 30 Si atoms) that the relaxation of the high strain induced by the curvature and dangling bond states at the surface is in direct competition with the preference of the interior atoms to be tetrahedrally coordinated. It was also observed that the type of core structure that forms depends on the mechanism by which the surface is passivated during the synthesis method.



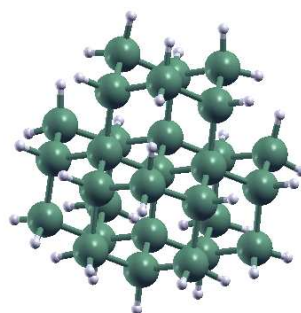
(c) Ge_5H_{12}



(d) $\text{Ge}_{10}\text{H}_{16}$



(e) $\text{Ge}_{14}\text{H}_{20}$



(f) $\text{Ge}_{29}\text{H}_{36}$

Figure 10.1: The Ge_nH_m nanoclusters.

10.3 Previous Experimental and Theoretical Work

10.3.1 Review of Experiments

There have been many experimental reports of strong visible photoluminescence (PL) in Ge nanocrystals in an SiO₂ matrix prepared by a variety of methods [75, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89]. In all of these reports the strong PL maximum was observed at about 2.2 eV, independent of size of the nanocrystals (in the 2 - 15 nm size range), and cannot be explained by a quantum confinement model. It has been suggested that defects in the SiO₂ matrix region or the interface between the SiO₂ matrix and the Ge nanocrystals are responsible for the size-independent visible PL [22].

Takeoka *et al.* [22] observed size-dependent PL for Ge nanocrystals in the 0.9 - 5.3 nm diameter size range in the near infrared region, which is closer to the bulk band gap of Ge (0.66 eV at 300 K), see Figure 10.2. The Ge nanocrystals were fabricated by rf cosputtering of Ge and SiO₂ and post annealing. For the sample with an average diameter d_{ave} of 5.3 nm, the PL peak was observed at about 0.88 eV. Decreasing the diameter of the nanocrystals had the effect of shifting the PL peak to higher energies, with the maximum of 1.54 eV for the sample with an $d_{ave} = 0.9$ nm. The PL intensity was also observed to increase with decreasing diameter.²

Takeoka *et al.* maintained that the strong PL detected in their experiments originated from a recombination of electron-hole pairs confined to the nanocrystal Ge [22]. However, Niquet *et al.* [58] arrived at a different interpretation, due to an incompatibility between their own semi-empirical (sp^3 tight binding) calculations and the experimental results. Niquet *et al.* [58] maintained that the size dependent PL in the near infrared region is probably related to the recombination of electron (hole) trapped on a surface defect with free hole (electron) confined in the nanocrystal. In addition, Niquet *et al.* [58] suggested that the size

²It is important to emphasize that size-dependent PL have been observed in Si nanoclusters [68, 229, 230, 231, 73, 232, 233, 74].

independent blue-green PL observed previously by so many researches, see Refs. [75, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89], comes from defects in the oxide (SiO_2 defects).

Taraschi *et al.* [54] optically characterized nanocrystalline Ge samples with infrared photoluminescence measurements at 4 K, using a 488 nm Ar laser excitation. The nanocrystalline Ge samples with small nanocrystals (<15 nm) exhibited a broad infrared PL background ranging from about $1.0 - 1.7 \mu\text{m}$. The broad background is attributed to quantum confinement in the Ge nanocrystals, consistent with a wide distribution of nanocrystal sizes and hence a broad range of confinement energies. The exact mechanism for the electron-hole recombination is not known: the possibilities include the recombination of both confined holes and electrons in nanocrystals or the recombination of an electron (hole) trapped on a nanocrystal surface defect, with a free hole (electron) confined in the nanocrystal.

More recently, Konchenko *et al.* [55] prepared spherical Ge dots of 3-10 nm in diameter on a 0.3 nm thick SiO_2 film of Si(111) substrate, similar to the procedure reported by Ichikawa *et al.* [234]. These nanocrystals can be reasonably regarded as a model of Ge nanoclusters embedded in an SiO_2 matrix. Quantum confinement of valence band of Ge nc observed by photoemission spectroscopy.

10.3.2 Theoretical Overview

Most of the theoretical work on Ge and Si nanoclusters has been dedicated to predicting the dependence of the optical gap with respect to the cluster size [23, 58, 25, 63, 64, 65, 24, 235, 26, 2, 56, 61, 1, 28]. For larger clusters (> 500 atoms), the optical gap can be approximated by calculating the quasiparticle gap and using *semi-empirical* approaches to correct the relatively small (few meV) exciton binding energy. However, in 0-2 nm size range these approaches break down due to poor descriptions of the exciton binding energy and an inadequate description of the nanocluster surface [2]. Our interest is is mainly confined to nanoclusters in this size range.

Semi-empirical treatments are based on knowledge of the electronic structure of the bulk

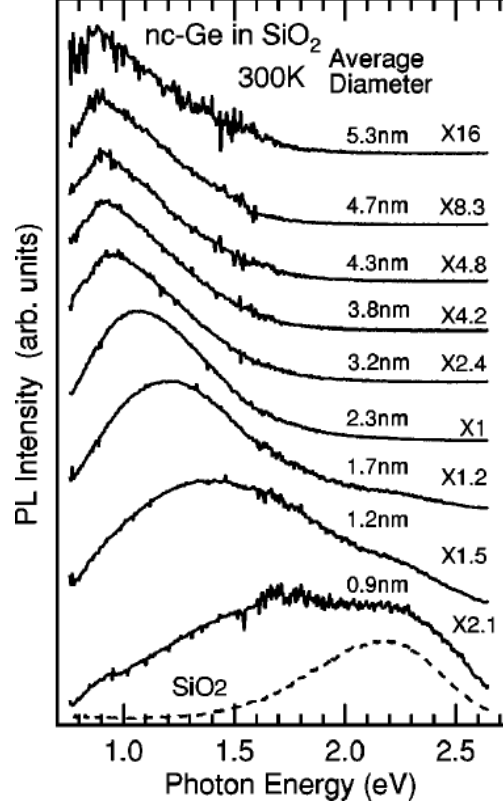


Figure 10.2: Experimental results for the dependence of photoluminescence spectra with respect to the average diameter of nano-crystalline Ge from Takeoka *et al.* [22]. The photoluminescence spectrum of the SiO₂ film is also displayed.

material (measured or computed) and the transferability to the nanoscale size range has to be questioned [58]. This holds for tight-binding and the empirical pseudopotential approach [63]. In addition, the tight-binding parametrization of bulk Ge is not unique and the extent to which this parametrization remains valid for small nanocrystallites is unknown [28]. The effective mass approximation is known to overestimate optical gaps for crystallites in the nanometer size region [236, 237]. Therefore, our interest is mainly confined to *ab-initio* work.

10.3.3 Focus on Ge Systems

There have been several studies of Ge nanoclusters by a variety of methods, including the effective mass approximation [238], tight binding [239, 58, 240], empirical pseudopotential

[59, 60] and *ab-initio* work such as DFT [58, 61, 62, 63, 64, 65], including TDDFT [1, 119, 28]. Some of the more notable work includes an sp^3 tight-binding study by Niquet *et al.* [58], who derived analytic laws for confinement energies over a whole range of sizes. Their work for the smaller nanoclusters was corroborated by *ab initio* LDA calculations. Melnikov and Chelikowsky [61] calculated the absorption spectrum in DFT-LDA using a real-space method. One of their most notable findings was that the spectra was essentially bulk-like for systems containing more than 250 atoms. In a related study, Melnikov and Chelikowsky [62] found that the calculated ionization potentials and affinities of the Ge_nH_m and the Si_nH_m nanoclusters remain significantly different from the corresponding bulk values up to systems containing around 800 atoms. Tsolakidis and Martin [1] and Nesher *et al.* [28] have calculated the absorption spectrum by TDDFT using the adiabatic local density approximation (ALDA). Weissker and co-workers have studied nanocluster Ge using a ΔSCF method ³ within LDA and LSDA [63, 65], where one study included structural relaxation of the ground state atomic configurations [64]. At the present time there are no published works that treat the optical excitations of the Ge nanoclusters in a many-body sense. The present QMC results are among the first truly many-body calculations of the excitations of the Ge nanoclusters.

10.3.4 Focus on Si Systems

Far more theoretical work has been dedicated to the study of Si nanoclusters. Among the many-body calculations for the Si_nH_m nanoclusters are the DMC results of Williamson *et al.* [2] in the 0 to 1.5 nm diameter size range, with the $\text{Si}_{87}\text{H}_{76}$ nanocluster being the largest in the study. Their DMC optical gaps were found to be in excellent agreement with earlier GW-BSE results by Rohlfing and Louie [25] for the first four fully hydrogenated Si_nH_m

³The ΔSCF has been briefly discussed in Section 2.3.2. Weissker *et al.* [63, 65, 64] used the ΔSCF method, which includes the electron-hole interaction energy, to estimate the optical gap by computing the total energy difference between the ground state and an excited state created by promoting an electron from the HOMO to the LUMO.

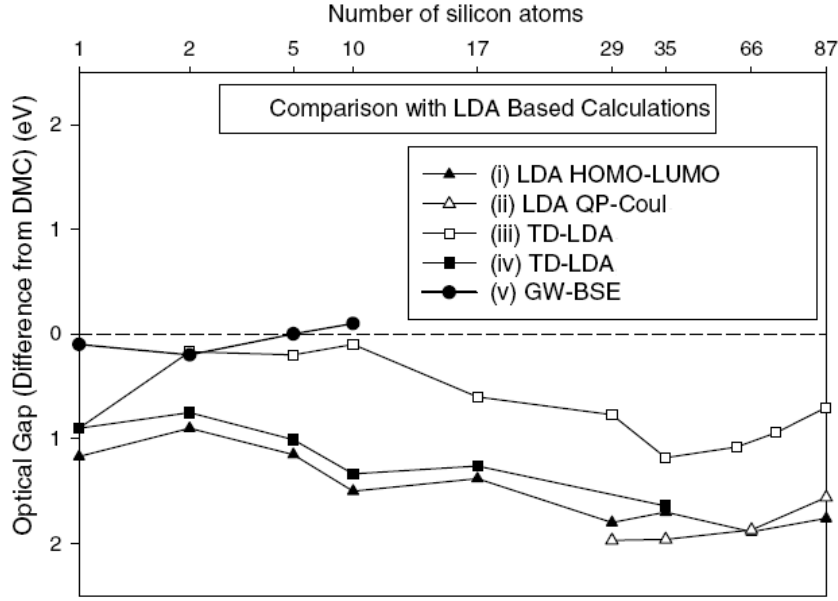


Figure 10.3: Difference in the optical gaps for the Si_nH_m nanoclusters predicted by several methods with respect to the DMC results of Williamson *et al.* [2]: \blacktriangle LDA HOMO-LUMO single particle gaps, \triangle LDA quasiparticle gap minus the empirical electron-hole exciton binding energy and corrected to include the electron-hole polarization energy from Ref. [23], \square TDLDA from Ref. [24] (effective gap) \blacksquare TDLDA, \bullet GW-BSE from Ref. [25].

nanoclusters: SiH_4 , Si_2H_6 , Si_5H_{12} and $\text{Si}_{10}\text{H}_{16}$ (Figure 10.3). However, it should be stated that the accuracy of these GW-BSE results[25] have since come into question because the authors neglected the off-diagonal matrix elements of the self-energy operator [57]. The inclusion of these terms shifted the lowest singlet of SiH_4 from 8.8 to 9.2 eV, which is closer to the DMC result of 9.1(1) eV. The full self-energy operator allows the unoccupied states to become more delocalized, which significantly reduces the electron-hole interaction [57].

A major focus of the work by Williamson *et al.* [2] was the comparison between the optical gaps predicted by DMC and DFT, including TDDFT. The DMC results for the optical gaps were found to be larger than the corresponding DFT-LDA single-particle gaps for all nanoclusters sizes. In addition, the DMC optical gaps are larger than the TDLDA gaps: this includes their own calculations and those of Vasiliev *et al.* [24]. The compari-

son between DMC and TDLDA is discussed in more detail in Section 10.6.1. Williamson *et al.* [2] also performed DFT calculations with the B3LYP functional [241]. The B3LYP results were found to be in good agreement with DMC, but for almost all nanocluster sizes the time-dependent B3LYP (TD-B3LYP) results were in worst agreement than the single-particle gaps: this is also true for the TD-B3LYP results of Garoufalis *et al.* [235]. In addition, Garoufalis *et al.* performed many-body multireference second-order perturbation theory (MR-MP2) calculations for Si nanoclusters as large as $\text{Si}_{29}\text{H}_{36}$. The MP2 gaps were found to nearly coincide with the TD-B3LYP gaps. Porter *et al.* [26] also performed DMC calculations on the Si_nH_m , $n + m \leq 34$, nanoclusters for the first dipole allowed and dipole forbidden singlet transitions and the first triplet transitions. The DMC results were compared to corresponding singles-only configuration interaction (CIS) and TDLDA results, see Figure 10.4, with the DMC optical gaps approximately 1 eV larger than the TDLDA gaps. This difference between the DMC and TDLDA gaps is similar to the difference independently reported by Williamson *et al.* [2].

Benedict *et al.* [56] compared the absorption spectra for several of the smaller Si_nH_m nanoclusters calculated by GW-BSE and TDLDA. In the GW-BSE calculations the screening is determined by the Hybertsen-Levine-Louie model [242, 132] and the optical gap is computed with DMC. The gross features of the two methods are similar for the larger Si_nH_m nanoclusters, due to the dominance of the Coulomb part of the interaction kernel which is the same in both treatments, but there are significant differences near the absorption edge. For the SiH_4 molecule, the two methods produce markedly different results, with the TDLDA spectra appearing closer to the experimental result.

10.3.5 The Ge - Si Crossover Question

A great deal of theoretical work has been dedicated to the Ge and Si crossover question: whether or not there is a crossover in the optical gaps of the Ge and Si nanocrystals at a critical radius (diameter). This is of great interest because it is known that in the bulk limit

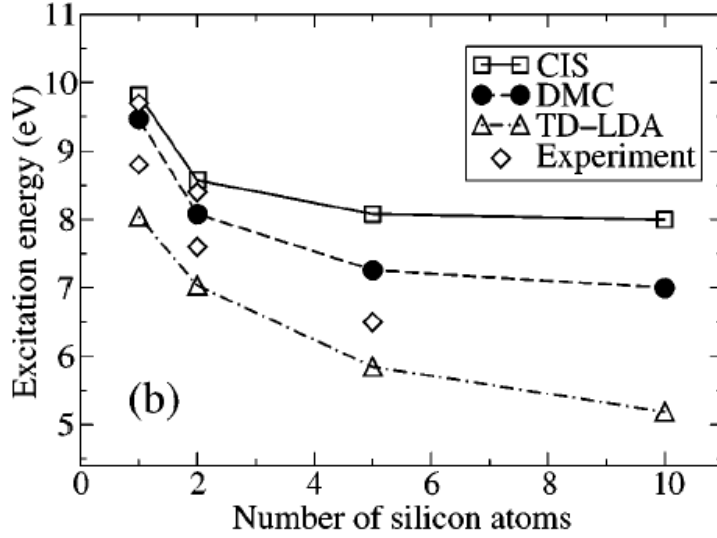


Figure 10.4: The lowest-bright (dipole-allowed) singlet excitation energies for several Si_nH_m nanoclusters by Porter *et al.* [26]. The experimental results for SiH_4 and Si_2H_6 are from Itoh *et al.* [19] and the results for Si_5H_{12} are from Fehér [27]. The first two peaks in the absorption spectrum are included for SiH_4 .

the band gap of Ge is approximately 0.68 eV, which is in the near infrared region (NIR), is lower than the band gap of bulk Si, which is approximately 1.1 eV.

Takagahara and Takeda [238], using an effective mass approximation (EMA), predicted the gaps of the Ge passivated nanoclusters would be larger than the Si nanoclusters for a critical radius smaller than 30 Å. This is attributed to differences between the electron and hole effective masses of Ge and Si, with the charge carriers of Ge having a smaller effective mass. Hill *et al.* [239], using tight binding (TB), came to a similar conclusion as Takagahara and Takeda [238], but with a critical radius less than 12 Å. Reboredo and Zunger [59], using an empirical pseudopotential approach, found no clear crossover for a radius as small as 10 Å. The empirical pseudopotential calculations indicated that the lowest conduction state changes from L derived to X derived below a critical diameter; hence the Ge nanocrystals become Si-like. Reboredo and Zunger [59] maintained that this phenomenon explained the

absence of a crossover. Weissker *et al.* in Refs. [63, 64, 65], using the Δ SCF method for DFT within the LDA and LSDA approximations, did not observe a crossover up to a radius of 12 Å. A recent paper by Nesher *et al.* [28] did not observe a crossover when comparing their TDLDA results for the absorption spectra for Ge_nH_m , ($n + m \leq 247$), to the corresponding TDLDA spectra for Si_nH_m by Vasiliev *et al.* [24]. The TDLDA optical gaps for Si_nH_m were larger than the Ge_nH_m optical gaps for all nanocluster sizes, including the optical gap of the “zero diameter” SiH_4 molecule (8.8 eV for SiH_4 compared to 8.3 eV for GeH_4). Both of these TDLDA studies calculated the absorption spectra within linear response theory, see Section 2.7.2 for a discussion. Real-time TDLDA calculations, see Section 2.7.3, by Tsolakidis and Martin [1, 119] also did not observe a crossover in the 1-2 nm diameter size range. In addition, *ab-initio* real-space pseudopotential calculations by Melnikov and Chelikowsky [62] found that while the ionization energies and electron affinities of the Si and Ge structures are nearly identical, the quasiparticle and single particle gaps of Si are larger by approximately 0.2 eV than those of Ge.

Closely related to the crossover question is the issue of the oscillator strengths: the relative size of the oscillator strengths for the Si and Ge clusters near the absorption edge. Tsolakidis and Martin [1] observed that the Ge nanoclusters exhibit more oscillator strength when compared to the Si nanoclusters in the low energy region near the absorption edge. This was attributed to the presence of more transitions in the low energy spectrum of bulk Ge than bulk Si [243]. Many of these transitions are indirect and therefore forbidden in the bulk, but because of the relaxation of \mathbf{k} selection rules due to quantum confinement are allowed in the nanoclusters. For both the Si and Ge nanoclusters it was found that the oscillator strengths of the indirect transitions are significantly enhanced. The results of Tsolakidis and Martin corroborate the *ab-initio* Δ SCF work within LDA by Weissker *et al.* [63], who observed that the Ge and Si nanoclusters behaved completely differently near the absorption edge (HOMO-LUMO gap). Weissker *et al.* [63] found that with increasing nanocluster size that the Si nanoclusters show a tail of weak transitions with decreasing

oscillator strengths, while optical transitions with large oscillator strengths occur for Ge nanoclusters near the HOMO-LUMO gap.

10.4 The Importance of Surface Chemistry

As we have already mentioned, quantum confined core states just represent one mechanism responsible for observed optical gap in the Ge and Si nanoclusters. The surface chemistry of the nanoclusters must be taken into account in order to properly explain their optical properties.⁴ Reducing the size of the clusters to the molecular limit results in an increase of the surface area to volume ratio; hence, the surface chemistry is expected to strongly influence the optical properties [66].

The most important mechanisms by which the surface chemistry affects the optical properties are by surface terminations, passivation and reconstruction. Termination refers to how the surface is terminated, specifically whether the surface is fully passivated or not. Passivation refers to the choice of passivates, such as single bonded passivates like hydrogen, or double bonded passivates like oxygen. Surface reconstruction refers to the structural relaxation of the nanocluster that occurs when surface passivants are removed.

TDLDA calculations by Vasiliev *et al.* [69] found that including a single oxygen atom on the surfaces of the $\text{Si}_{35}\text{H}_{34}$ and $\text{Si}_{29}\text{H}_{34}$ nanoclusters (to create $\text{Si}_{35}\text{H}_{34}\text{O}$ and $\text{Si}_{29}\text{H}_{34}\text{O}$) effectively reduced the optical gap by 1-2 eV. It is believed that oxygen introduces extra localized states near the gap [69]. This work was motivated by the experimental results of Wolkin *et al.* [68], who observed a large redshift of the photoluminescence in porous silicon after exposure to air. More specifically, a redshift of the order of 1 eV was reported for the samples composed of crystallites smaller than 2 nm which was attributed to surface oxidation of silicon nanocrystals.

DFT and DMC calculations by Puzder *et al.* [66] showed that significant changes occur

⁴The work cited in this section deals with Si nanoclusters; however, it is reasonable to expect that any conclusions can be extended to Ge nanoclusters

in the gap of fully hydrogenated nanoclusters when the surface contains passivants other than hydrogen. Substituting hydrogen for another single bonded passivant was found to have an insignificant affect on the optical properties of the $\text{Si}_{35}\text{H}_{36}$ nanocluster. In contrast, double bonded passivants ($\text{Si}_{35}\text{H}_{34}\text{Y}$) reduced the gap by at least 1 eV, this reduction was accompanied by a local distortion of the sp^3 network on the surface. It was proposed that the origin of the wide discrepancy between previous theoretical results [24, 236, 244] for hydrogen passivated Si nanoclusters and experimental results in Refs. [227, 232, 245] could be due to oxygen contamination.

Mitas *et al.* [67] measured the effect of surface reconstruction for the $\text{Si}_{29}\text{H}_{36}$ nanocluster: the absorption spectrum was determined by CIS and subsequently shifted such that the CIS absorption edge (lowest excitation energy) was in agreement with the optical gap determined by DMC. Mitas *et al.* [67] found that removing 12 hydrogen atoms from the fully passivated $\text{Si}_{29}\text{H}_{36}$ nanocluster and allowing for surface reconstruction lowered the absorption edge from approximately 4.8 eV to approximately 3.5 eV. Removing an additional 12 hydrogen atoms (to form $\text{Si}_{29}\text{H}_{12}$) and allowing for reconstruction dramatically lowered the absorption edge to approximately 0.1 eV. In addition, a Density Functional study Lehtonen and Sundholm [70] found that the absorption threshold of the fully passivated $\text{Si}_{29}\text{H}_{36}$ nanocluster lies in the ultraviolet (320-340 nm) 3.87-3.65 eV region, whereas the absorption threshold of the $\text{Si}_{29}\text{H}_{24}$ nanocluster is about 450 nm or 2.76 eV.

10.5 Computational Method

The computation details regarding the QMC trial wave function of the Ge_nH_m nanoclusters are identical to the description in Section 9.2, for the GeH_4 and Ge_2H_6 molecules. All DMC calculations use a time-step of either 0.05 a.u. or 0.025 a.u.

The optical excitations are calculated in QMC following the method in Section 4.13. Only the trial wave function for the ground-state is optimized; optimization of the ex-

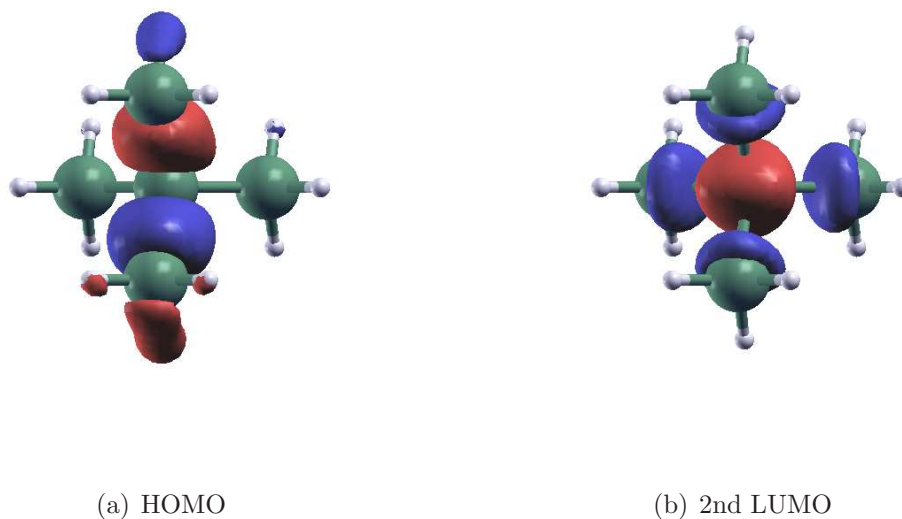


Figure 10.5: Ge_5H_{12} : Isosurfaces for the molecular orbitals used to compute the optical gap (red +, blue -).

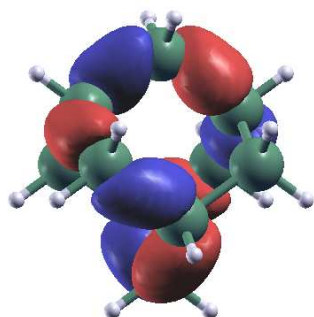
cited state wave function was found to have an insignificant effect on the optical gap. Figures 10.5, 10.6, 10.7 and 10.8 plot the isosurfaces of the HOMO and excited state MOs used in computing the optical gaps of Ge_5H_{12} , $\text{Ge}_{10}\text{H}_{16}$, $\text{Ge}_{14}\text{H}_{20}$ and $\text{Ge}_{29}\text{H}_{36}$.⁵ The isosurfaces of the MOs for GeH_4 and Ge_2H_6 can be found in Figures 9.3 and 9.4. Note that the second LUMO is used to compute the optical gap of Ge_5H_{12} , since the $\text{HOMO} \rightarrow \text{LUMO}$ transition is optically forbidden.

For the larger nanoclusters, it was necessary to constrain the weights in the correlated sampling to unity for stability reasons in the optimization [192, 51]. For a discussion of constrained weight optimization see Section 6.8.4.

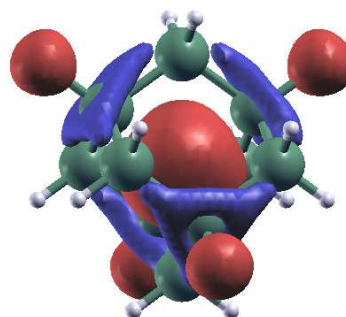
10.5.1 The LDA Relaxation of the Structures

The LDA relaxation of the structures closely follows the method of Tsoiakidis and Martin in Refs. [1, 119]. The structural relaxations include all of the constituent atoms; all calcu-

⁵All isosurfaces made using the XCrySDen package [220].

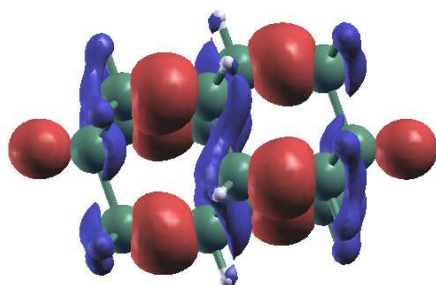


(a) HOMO

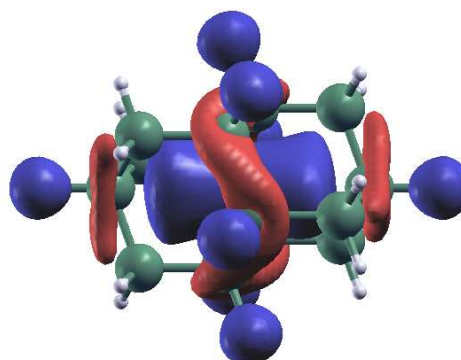


(b) LUMO

Figure 10.6: $\text{Ge}_{10}\text{H}_{16}$: Isosurfaces for the molecular orbitals used to compute the optical gap.



(a) HOMO



(b) LUMO

Figure 10.7: $\text{Ge}_{14}\text{H}_{20}$: Isosurfaces for the molecular orbitals used to compute the optical gap.

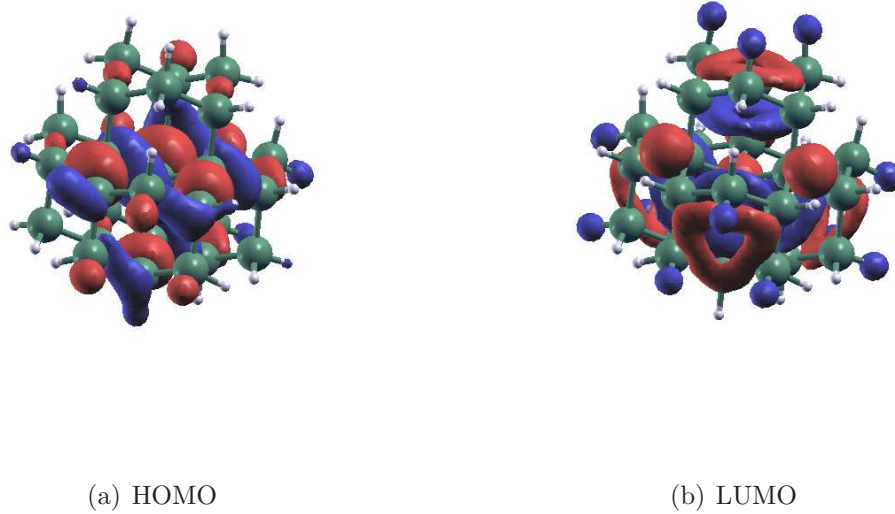


Figure 10.8: $\text{Ge}_{29}\text{H}_{36}$: Isosurfaces for the molecular orbitals used to compute the optical gap.

lations were performed using the SIESTA DFT package [181, 246, 247, 248]. The SIESTA calculations featured an auxiliary real space cutoff that is equivalent to a plane-wave cutoff of 70 Ry.

All of the SIESTA calculations used a relativistic LDA pseudopotential, generated by the Troullier-Martins [206] method. The choice of pseudopotential is one of the most important differences between the results presented in this thesis and those of Tsolakidis and Martin and the majority of the previous work on Ge clusters. In addition, in SIESTA it is necessary to use a pseudopotential for the H nuclei: this is because SIESTA cannot deal with the cusp at the location of the H nuclei.

The basis for each Ge atom consists of 13 numerical atomic orbitals (NAO's) per atom: two radial shapes to represent the two $4s$ states with confinement radii $r_s = 5.08$ a.u., two radial shapes to represent the six $4p$ states with confinement radii $r_p = 8.04$ a.u. and a polarization d shell with confinement radius $r_d^{pol} = 8.04$ a.u. The H basis set includes five NAO's per atom: two radial shapes for $1s$ states with confinement radii $r_s = 6.05$ a.u. and a p polarization shell with a confinement radius of $r_p = 6.05$ a.u. In SIESTA this is the

default DZP basis for both Ge and H. The basis for Si is very similar to the Ge basis, with the confinement radii $r_s = 6.12$ a.u., $r_p = 7.85$ a.u., and $r_d^{pol} = 7.85$ a.u. for the polarization d states.

10.6 Results for the Optical Gaps

10.6.1 Comparison of DMC to TDLDA for Ge Clusters

The TDLDA spectrum has generally been observed to evolve from a discrete spectrum with large oscillator strengths near the absorption edge for the smallest Ge_nH_m and Si_nH_m nanoclusters to a quasicontinuous spectrum with small oscillator strengths for the larger nanoclusters [56, 24, 28]. As we will see this will have a consequence when determining the optical gaps of the larger nanoclusters. All of the nanoclusters under consideration are well below the bulk limit where the TDLDA results are expected to converge to the LDA results [118].

It should also be stated that it has been observed by several researches that TDLDA does not dramatically correct the value of the optical gap of the Si and Ge nanoclusters over LDA; however, TDLDA gives better results for the absorption spectrum over the entire energy range [2, 56, 1, 119, 249]. This is due to the inclusion of dynamical correlations within TDLDA which corrects the oscillator strengths of the transitions and moves spectrum to higher energies (blueshifts) compared to the LDA spectrum. Tsolakidis and Martin observed that TDLDA corrected the gap by only about 0.2 eV relative to LDA [1].

From Figure 10.9 it can be seen that the DMC+CPP gaps for the Ge_nH_m nanoclusters are larger than the corresponding TDLDA gaps reported by Nesher *et al.* [28], which is in agreement with the previous DMC and TDLDA comparisons made for the Si_nH_m nanoclusters by Williamson *et al.* [2] (see Figure 10.3) and Porter *et al.* [26] (see Figure 10.4). Note that the TDLDA gaps shown in the figure are “effective” optical gaps, which are described below in the discussion about silicon clusters. Tsolakidis and Martin [1] have also performed

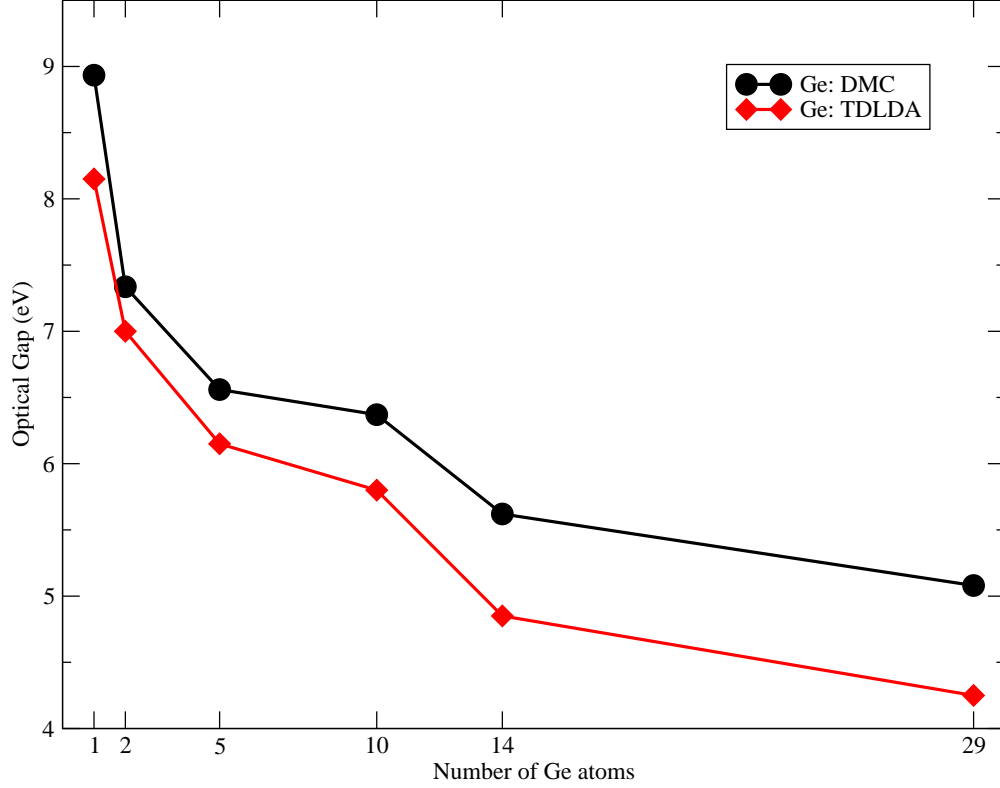


Figure 10.9: Comparison of the present DMC+CPP results for the lowest gaps of the Ge_nH_m , $(n + m) \leq 65$, nanoclusters (\bullet) to the TDLDA results of Nesher *et al.* [28] (\blacklozenge). The TDLDA results are “effective gaps” as described in the text.

TDLDA calculations for the Ge_nH_m nanoclusters, with $\text{Ge}_{29}\text{H}_{36}$ as their smallest structure. Their value for the lowest gap of 3.66 eV for $\text{Ge}_{29}\text{H}_{36}$ [119] is approximately 0.5 eV lower than the effective optical gap reported by Nesher *et al.* [28], which is also consistent with the comparison of the lowest and effective gaps calculated using TDLDA for silicon clusters [2]. See Figure 10.10 for the TDLDA absorption spectra from both Nesher *et al.* [28] and Tsolakidis and Martin [1].

The size of the difference between the optical gaps from DMC and TDDFT strongly depends on how the optical gap is assigned in TDDFT. Williamson *et al.* [2] choose to compare their DMC results for the Si_nH_m , $(n + m) \leq 163$, nanoclusters to the first non-zero transition in the TDDFT spectrum, maintaining that this is the closest comparison between their DMC results and TDLDA. Their DMC results for the optical gaps were found to be

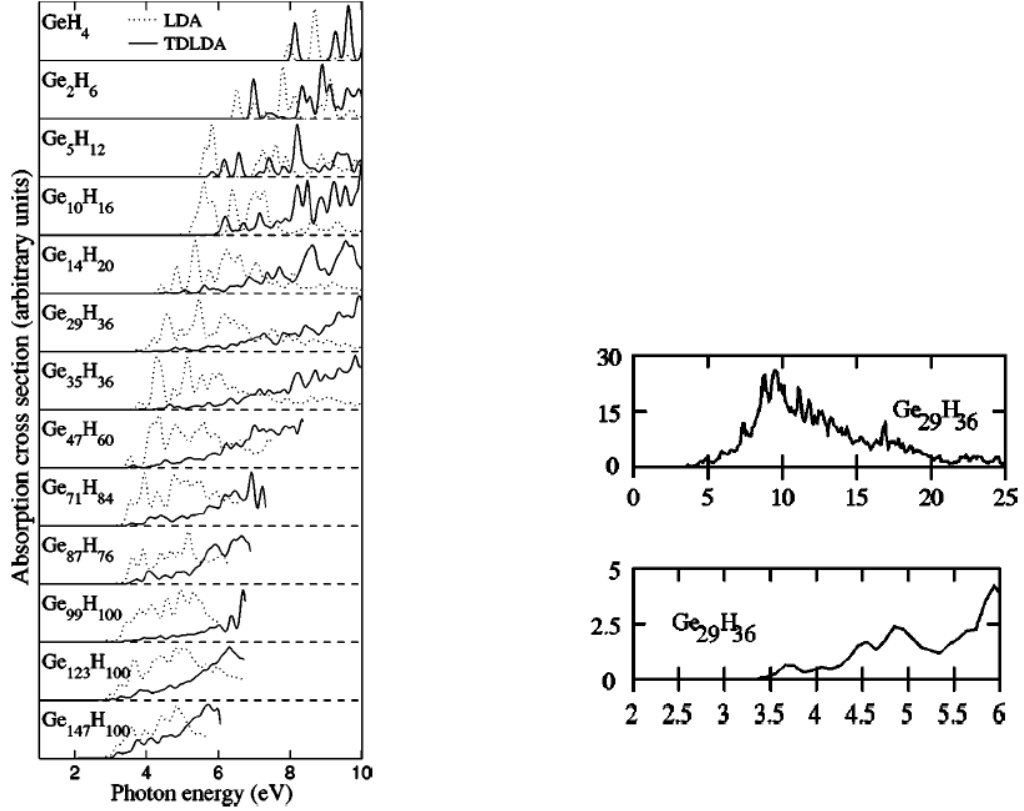


Figure 10.10: Left: TDLDA results for the absorption spectra of the Ge_nH_m nanoclusters by Nesher *et al.* [28]. Right: TDLDA results for the absorption spectrum of $\text{Ge}_{29}\text{H}_{36}$ by Tsolakidis and Martin [1], plotted on two energy scales.

larger by 1-2 eV than the TDLDA gaps for all nanocluster sizes; the size of the difference was smaller when comparing to the “effective” optical gaps from the TDLDA results of Vasiliev *et al.* [24]. Vasiliev *et al.* [24] and Nesher *et al.* [28] defined the TDLDA optical gap as the point at which the integrated oscillator strength exceeds a threshold of 2% of the total absorption spectrum. The argument behind this assignment is that the oscillator strengths of the transitions below the integration threshold would be below the experimentally detectable limit [24, 28]. For the smallest nanoclusters these different methods of assigning the optical gap are in agreement; the oscillator strengths of the first excitations exceeds the threshold. A difference between the two approaches begins to appear for the larger nanoclusters; the integration threshold is exceeded only after integrating over many smaller peaks in the

absorption spectrum.

10.6.2 Comparison to Previous DMC Results for Si Clusters

From Figure 10.11 it is immediately noticeable that a difference exists between the two sets of previous DMC calculations for the optical gaps of the Si_nH_m nanoclusters. This is not a disagreement, but rather the fact that Porter *et al.* [26] have given the gap for a higher transition whereas the other QMC results are for the lowest gap. For the $\text{Si}_{10}\text{H}_{16}$ nanocluster Porter *et al.* [26] choose to assign the third dipole allowed transition from CIS to calculate the optical gap in DMC. The justification for this choice was that the oscillator strengths of the first two transitions were considerably smaller; at the CIS level the difference between the first transition and the third transition was about 0.7 eV.

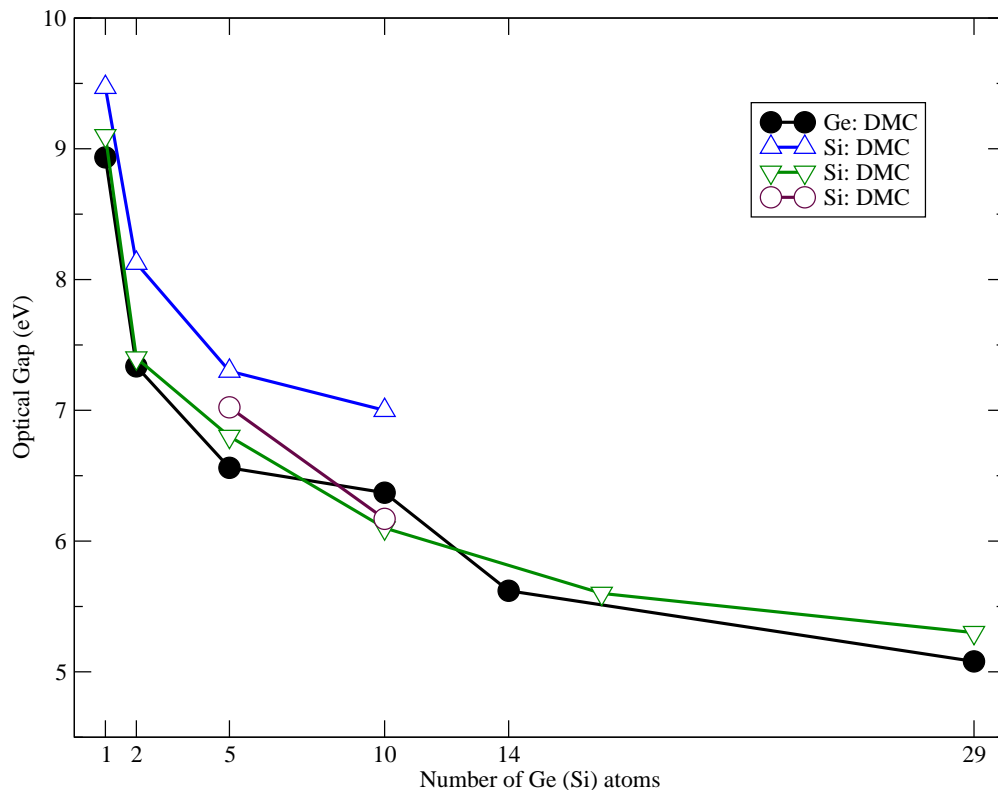


Figure 10.11: Comparison of our DMC results for Ge_nH_m (●) to the DMC results for the corresponding Si nanoclusters by Porter *et al.* [26] (Δ) and Williamson *et al.* [2] (▽). We have also included our own DMC results (○) for Si_5H_{12} and $\text{Si}_{10}\text{H}_{16}$.

Figure 10.11 also shows results of our own DMC calculations for the optical gaps of Si_5H_{12} and $\text{Si}_{10}\text{H}_{16}$ in order to have a direct comparison of the Si and Ge clusters, and to compare with the previous work on Si clusters. We have not included the CPP for Si, which is a much smaller effect in Si than in Ge and cannot be approximated by a local potential [43]. Otherwise, all of the computational details, including the use of a HF pseudopotential [12] and a HF derived trial wave function, are identical for the Si_nH_m and Ge_nH_m nanoclusters. Our DMC gap for Si_5H_{12} lies in between the DMC results of Williamson *et al.* [2] and Porter *et al.* [26], while our DMC gap for $\text{Si}_{10}\text{H}_{16}$ is nearly identical to the result of Williamson *et al.* [2].

It is gratifying that the lowest gaps calculated by QMC are in good agreement, even though there are differences in the calculations. Williamson *et al.* [2] used LDA pseudopotentials to replace the Si and H cores and DFT-LDA single particle orbitals to construct their Slater determinants. The single particle orbitals were first transformed by a unitary transformation to form a set of truncated, maximally localized Wannier (MLW) functions [250, 251], which have been shown to increase the efficiency of QMC.⁶ In contrast, our work and the calculations of Porter *et al.* [26] used a relativistic HF pseudopotential for Si. To obtain the Slater determinant part of the trial wave function, Porter *et al.* [26] used the results of a CIS calculation. Porter *et al.* [48, 26] argued that the CIS derived trial wave function was considerably better than the HF derived wave functions for the excited states based on tests for the SiH_4 molecule. It is important to note that CIS, which involves only single-substitutions, differs from HF for the excited states; the ground state for CIS is the HF ground state.

Table 10.1 contains all of our QMC results for the optical gaps of the Ge_nH_m , $(n + m) \leq 65$. This includes the VMC+CPP and DMC+CPP gaps along with the CPP corrections as calculated by Eq. (9.2). This table serves as a reference for future comparisons.

We have not been able to observe any clear size dependence for the CPP correction to the

⁶Recent articles have claimed near linear scaling in QMC by use of MLW functions [252, 253].

Table 10.1: QMC optical gaps of the Ge_nH_m , $(n + m) \leq 65$, nanoclusters with the CPP correction ΔCPP . All energies are in eV.

Cluster	VMC+CPP	ΔCPP	DMC+CPP	ΔCPP
GeH_4	9.7651(74)	-0.2520(14)	8.934(13)	-0.2272(21)
Ge_2H_6	7.707(11)	-0.2357(25)	7.336(9)	-0.2235(16)
Ge_5H_{12}	6.845(18)	-0.318(3)	6.562(31)	-0.3045(46)
$\text{Ge}_{10}\text{H}_{16}$	6.436(27)	-0.074(3)	6.36(18)	-0.0678(89)
$\text{Ge}_{14}\text{H}_{20}$	5.495(24)	-0.038(3)	5.62(12)	-0.033(14)
$\text{Ge}_{29}\text{H}_{36}$	5.480(73)	-0.133(16)	5.09(21)	-0.11(2)

optical gaps. We believe that the CPP correction strongly depends on the properties of the excited state used to calculate the optical gap in DMC. For the larger Ge_nH_m nanoclusters, the size of the CPP corrections are roughly the same magnitude as the error bars of the optical gaps. This is not surprising since in the bulk limit the CPP correction for the indirect band gap of Ge is estimated to be approximately 0.2 eV. This estimate is based on quasiparticle band energies provided by Eric Shirley [20]. For the largest nanocluster in our study, $\text{Ge}_{29}\text{H}_{36}$, the CPP correction is approximately 0.1 eV while the error bar for the optical gap is approximately 0.2 eV. In addition, it should be stated that it becomes impractical to reduce the error bars of the energy differences below 0.2 eV for the larger nanoclusters due to the lack of experimental results for comparison. Furthermore, systematic errors in DMC due to the fixed-node error the evaluation of the non-local pseudopotentials within the locality approximation limit the accuracy of the optical gaps and obviate reducing the error bars to the level of chemical accuracy (approximately 0.04 eV).

Chapter 11

Conclusions, General Observations and Future Directions

This work presented in this thesis is related to four topics: the treatment of electronic correlations among valence electrons, inclusion of the effects of the core on the valence states, applications of Quantum Monte Carlo (QMC) and optical properties of semiconductor nanoclusters. The highly accurate QMC method can treat electronic correlations at a many-body level; however, the large energies of core states leads to large fluctuations and it almost always necessary to remove the core-electrons of heavy elements and treat the valence electrons using a pseudopotential. Therefore the core-valence interactions are treated at a different level of accuracy than the valence-valence interactions. The post transition element Ge ($Z=32$) is of particular interest since the usual method of treating the core-valence interactions by a pseudopotential from an independent particle approximation is unsatisfactory. An improved method is to augment a Hartree-Fock (HF) pseudopotential with a core polarization potential (CPP) to take into account core-valence correlations at a many-body level, which are large in Ge because of its shallow, polarizable core.

An example of a class of systems where electronic correlations are known to be important are semiconductor nanoclusters. The optical properties of nanoclusters constructed from Ge or Si have recently been the focus of a growing body of experimental and theoretical work. These nanoclusters have a widened band gap relative to the bulk due to quantum confinement that is inversely proportional to the size (diameter) of the nanoclusters. Therefore, Ge and Si nanoclusters offer the possibility of tunable photoluminescence which has a tremendous number of potential applications. This thesis is one of the earliest applications of QMC to the optical excitations of Ge nanoclusters.

QMC calculations were presented for the ground state and excited states of the Ge atom, molecules and hydrogen terminated Ge nanoclusters, with $\text{Ge}_{29}\text{H}_{36}$ as the largest system in the study. As mentioned previously, the interacting system of valence electrons was explicitly treated at the many-body QMC level using a Hartree-Fock pseudopotential to which is added a CPP to account for the core-valence correlations. In addition, the effect of core-valence correlation can be quantified by the estimator of the CPP. The CPP correction to excitations can be evaluated as the total energy difference of the ground state and excited state CPP estimators.

The effect of the CPP was found to be significant for the total energy and excitations of the atom and small molecules. Analysis of these cases shows that size of the CPP correction for the excitations largely depended on the nature of the excited states; mainly the relative Ge s and p character of the excited state. The CPP correction is largest for the excited states with a dominant Ge s character. For the larger Ge_nH_m nanoclusters it was found that the size of the CPP corrections were much smaller than in the Ge atom and the molecules GeH_4 and Ge_2H_6 , with no clear size dependence. This can be understood due to the strong mixing of s and p Ge atomic orbitals for the excited states.

The QMC optical gaps for the Ge_nH_m nanoclusters are compared to the optical gaps determined from time-dependent density functional theory (TDDFT) in the adiabatic local density approximation (TDLDA) [28, 1] and to QMC results for the corresponding Si nanoclusters [2, 26]. The calculations were carried out using structures optimized with DFT calculations in the Local Density Approximation (LDA) and using a Gaussian basis set for the trial wave functions in QMC. Optical excitations were treated by promoting one electron from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO) orbitals, as has been done in previous work [57, 2]. The QMC results for the gaps were found to be larger than the TDLDA gaps by roughly one electron volt, which is in agreement with previous comparisons between QMC and TDLDA for Si_nH_m in the 0 - 2 nm diameter size range [26, 2]. In order to make direct comparisons of Ge and Si clusters

it was necessary to perform calculations of Si (Si_5H_{12} and $\text{Si}_{10}\text{H}_{16}$) using the same methods, with results that agree well with previous calculations. The primary conclusion is that the gaps for Ge clusters are very close to those of Si clusters, as shown in Figure 10.11. It is worthwhile to note that previous TDLDA [1] and delta self-consistent field theory within DFT-LDA [63] calculations found that oscillator strengths are significantly higher for Ge than for Si clusters. Thus the work in this thesis supports the conclusion that the gaps of Si and Ge clusters are very close, but nevertheless Ge clusters have increased optical absorption at low energies due to increased oscillator strengths.

A major part of the work for this thesis was dedicated to the design and development of a new Quantum Monte Carlo program named qmcPACK [3]: this includes writing the actual code and extensive testing. This code can perform Variational Monte Carlo and Diffusion Monte Carlo in the Fixed-Node approximation; the completion of these two components was a prerequisite for the results in the preceding chapters. A factor in the design of qmcPACK was versatility in the choice of the trial wave function. qmcPACK can import trial wave functions from a variety of sources, such as the SIESTA DFT package, Gaussian 98/03, Ab-Init and orbitals on radial or three-dimensional grids. In addition, qmcPACK can now handle periodic boundary conditions and can do twist-averaging boundary conditions [254]. The long-ranged electron-electron and electron-ion interactions can either be handled by the traditional Ewald break-up method ¹ or by the method developed by Natoli and Ceperley [255] and modified by Esler [256].

In the future it is desirable to use better trial wave functions, especially for excited states. The most common approach to construct an excited state in QMC is to promote an electron from an occupied to an unoccupied excited state orbital in the Slater determinant part: the single-particle orbitals usually come from a preceding Hartree-Fock or DFT calculations. It is well known that the excited states from both HF and DFT in the standard approximations such as LDA suffer from deficiencies. One possible improvement is to use multiple deter-

¹See Appendix F of Ref. [29] for a discussion.

minant wave functions from Configuration Interaction or Coupled Cluster approximations, such approaches have been used in Refs. [48, 26, 67]. Another possibility is the correlated geminal approach discussed in Ref. [190]. A minor modification is to use the Padé-Jastrow (see Section 6.6.1) for the hydrogen-electron correlations; the present work uses a cusplless function (see Section 6.6.2). The purpose of using the Pade-Jastrow is to better satisfy the cusp condition that occurs at the location the hydrogen nuclei, which is only approximately satisfied by the Slater type orbitals for the hydrogen $1s$ and $2s$ orbitals. The Jastrow parameter a can be optimized as opposed to being explicitly set to satisfy the cusp condition: $a = Z$, where Z is the charge of ion core.

Possible future projects include simulating larger Ge_nH_m nanoclusters and performing calculations on bulk Ge. The results for the gaps of the nanoclusters can be compared to TDLDA results in Refs. [1, 28] and QMC results for Si nanoclusters in Ref. [2]. This would be the most obvious extension of the current work. For the solid, QMC can be used to calculate the band structure of Ge; of particular interest is the effect of the CPP on the band gaps, especially the direct and indirect band gaps of Ge. Previous QMC studies of Ge have reported band gaps at a few special k -points, but the details of these calculations regarding the choice of pseudopotential and trial wave function are much different from the approach outlined in this thesis [49, 50]. These calculations were an early application of QMC to periodic systems and were therefore intended to establish the viability of QMC for solids, rather than highly accurate results.

Appendix A

Appendix

A.1 The Spherical Harmonics

The spherical harmonics are given by the expression

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos \theta) e^{im\phi} \quad (\text{A.1})$$

where P_l^m is an associated Legendre function defined by

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} (1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l. \quad (\text{A.2})$$

In order for P_l^m to be finite on the interval $-1 \leq x \leq 1$ the parameter l must be a zero or a positive integer and the integer m is restricted to the values $-l, -(l-1), \dots, 0, \dots, (l-1), l$. Eq. (A.2) is valid for both positive and negative m : it can be shown that

$$P_l^{-m} = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x); \quad (\text{A.3})$$

hence the spherical harmonics obey the following relation

$$Y_{l,-m}(\theta, \phi) = (-1)^m [Y_{lm}(\theta, \phi)]^*. \quad (\text{A.4})$$

The spherical harmonics obey the normalization and orthogonality conditions over the

sphere

$$\int d\Omega [Y_{lm}(\theta, \phi)]^* Y_{l'm'}(\theta, \phi) = \delta_{ll'} \delta_{mm'}, \quad (\text{A.5})$$

where

$$\int d\Omega = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta$$

is the differential area element on the surface of the sphere. Another important relation obeyed by the spherical harmonics is the completeness relation

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l [Y_{lm}(\theta', \phi')]^* Y_{lm}(\theta, \phi) = \delta(\phi - \phi') \delta(\cos\theta - \cos\theta'), \quad (\text{A.6})$$

which in the more common Dirac notation is

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l |Y_{lm}\rangle \langle Y_{lm}| = 1. \quad (\text{A.7})$$

Therefore, it is possible to express any function of θ and ϕ , (on the surface of a sphere) in terms of the complete set of spherical harmonics

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\theta, \phi). \quad (\text{A.8})$$

with the coefficients c_{lm} given by

$$c_{lm} = \int d\Omega f(\theta, \phi) [Y_{lm}(\theta, \phi)]^* \quad (\text{A.9})$$

Another important identity is the summation rule

$$\sum_{m=-l}^l |Y_{lm}(\theta, \phi)|^2 = \frac{2l+1}{4\pi}; \quad (\text{A.10})$$

this identity has important implications for closed-shell atomic systems. It should be noted

that this section largely follows the formalism of Jackson [111] and that different conventions exist for dealing with the phase factor, $(-1)^m$, of the spherical harmonics [180]. Table A.1 contains the spherical harmonics up to $l = 3$.

Table A.1: The spherical harmonics $Y_{lm}(\theta, \phi)$ for the s, p, d and f channels.

l	l, m	Y_{lm}
s	$0, 0$	$\sqrt{\frac{1}{4\pi}}$
p	$1, \pm 1$	$\mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi}$
	$1, 0$	$\sqrt{\frac{3}{4\pi}} \cos \theta$
d	$2, \pm 2$	$\sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}$
	$2, \pm 1$	$\mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi}$
	$2, 0$	$\sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)$
f	$3, \pm 3$	$\mp \sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{\pm 3i\phi}$
	$3, \pm 2$	$\sqrt{\frac{105}{32\pi}} \sin^2 \theta \cos \theta e^{\pm 2i\phi}$
	$3, \pm 1$	$\mp \sqrt{\frac{21}{64\pi}} \sin \theta (5 \cos^2 \theta - 1) e^{\pm i\phi}$
	$3, 0$	$\sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta)$

A.1.1 The Real Spherical Harmonics

The spherical harmonics, see Appendix A.1, are complex functions except for the components $m = 0$, *i.e.* when the z -component of the angular momentum is equal to zero. We would like to avoid using complex functions in electronic structure codes and determine a real counterpart to the spherical harmonics, *i.e.* $Y_{lm} \rightarrow S_{lm}$. The convention for the real spherical harmonics used in the Gaussian03 [11] package:

$$S_{lm} = \begin{cases} (-1)^m \sqrt{2} \Re(Y_l^{|m|}) & m > 0 \\ Y_l^0 & m = 0 \\ (-1)^m \sqrt{2} \Im(Y_l^{|m|}) & m < 0 \end{cases}, \quad (\text{A.11})$$

where the factor $\sqrt{2}$ is necessary for normalization. Table A.2 contains the real spherical harmonics for the s, p, d and f angular momentum channels. The functions can be transformed from the Cartesian representation to spherical coordinates by the following change of variables

$$x = r \sin(\theta) \cos(\phi) \quad (\text{A.12})$$

$$y = r \sin(\theta) \sin(\phi) \quad (\text{A.13})$$

$$z = r \cos(\theta). \quad (\text{A.14})$$

Table A.2: The real spherical harmonics $S_{lm}(x, y, z)$ in Cartesian coordinates for the s, p, d and f channels. The second column is the index $I_{lm} = l(l+1) + m$.

	l, m	I_{lm}	$r^l S_{lm}$	$\nabla(r^l S_{lm})$
s	0,0	0	$\sqrt{\frac{1}{4\pi}}$	$\{0, 0, 0\}$
p	1,-1	1	$\sqrt{\frac{3}{4\pi}}y$	$\sqrt{\frac{3}{4\pi}}\{0, 1, 0\}$
	1,0	2	$\sqrt{\frac{3}{4\pi}}z$	$\sqrt{\frac{3}{4\pi}}\{0, 0, 1\}$
	1,1	3	$\sqrt{\frac{3}{4\pi}}x$	$\sqrt{\frac{3}{4\pi}}\{1, 0, 0\}$
d	2,-2	4	$\sqrt{\frac{15}{4\pi}}xy$	$\sqrt{\frac{15}{4\pi}}\{1, 1, 0\}$
	2,-1	5	$\sqrt{\frac{15}{4\pi}}yz$	$\sqrt{\frac{15}{4\pi}}\{0, 1, 1\}$
	2,0	6	$\sqrt{\frac{5}{16\pi}}(3z^2 - r^2)$	$\sqrt{\frac{5}{4\pi}}\{-x, -y, 2z\}$
	2,1	7	$\sqrt{\frac{15}{4\pi}}xz$	$\sqrt{\frac{15}{4\pi}}\{1, 0, 1\}$
	2,2	8	$\sqrt{\frac{15}{16\pi}}(x^2 - y^2)$	$\sqrt{\frac{15}{4\pi}}\{1, 1, 0\}$
f	3,-3	9	$\sqrt{\frac{35}{32\pi}}(3x^2 - y^2)y$	$\sqrt{\frac{35}{32\pi}}\{3x^2 - 3y^2, -6xy, 0\}$
	3,-2	10	$\sqrt{\frac{105}{4\pi}}xyz$	$\sqrt{\frac{105}{4\pi}}\{yz, xz, xy\}$
	3,-1	11	$\sqrt{\frac{21}{32\pi}}(5z^2 - r^2)y$	$\sqrt{\frac{21}{32\pi}}\{2xz, -2yz, x^2 - y^2\}$
	3,0	12	$\sqrt{\frac{7}{16\pi}}(5z^2 - 3r^2)z$	$\sqrt{\frac{7}{16\pi}}\{-6xz, -6yz, 6z^2 - 3x^2 - 3y^2\}$
	3,1	13	$\sqrt{\frac{21}{32\pi}}(5z^2 - r^2)x$	$\sqrt{\frac{21}{32\pi}}\{4z^2 - 3x^2 - y^2, -2xy, 8xz\}$
	3,2	14	$\sqrt{\frac{105}{16\pi}}(x^2 - y^2)z$	$\sqrt{\frac{105}{16\pi}}\{2xz, -2yz, x^2 - y^2\}$
	3,3	15	$\sqrt{\frac{35}{32\pi}}(x^2 - 3y^2)x$	$\sqrt{\frac{35}{32\pi}}\{3x^2 - 3y^2, -6xy, 0\}$

A.2 Radial Grids

This section contains the analytic forms of the three types of radial grids used in this thesis. Each of the grids has N points and the index i runs from 0 to $N - 1$. The most basic grid is the linear grid. The logarithmic grids are desirable for applications that require a large number of points near the origin, such as atoms with the $-Z/r$ Coulomb potential. The initial point of the logarithmic grid r_0 cannot be zero and is typically chosen to be a very small number. In contrast the initial point of the zero based logarithmic grid is zero.

A. The linear grid:

$$r_i = \left(\frac{r_f - r_0}{N - 1} \right) i + r_0 \quad (\text{A.15})$$

B. The logarithmic grid:

$$r_i = r_0 \left(\frac{r_f}{r_0} \right)^{\frac{i}{N-1}} \quad (\text{A.16})$$

C. The zero based logarithmic grid:

$$r_i = B (\exp(Ai) - 1) \quad (\text{A.17})$$

A.3 One-Dimensional Cubic Splines

Given a function defined on a grid, $\{x_i\}, i = 1 \dots N$, such that $y_i = y(x_i)$, we would like to evaluate the function (interpolate) at a point x in the interval $[x_j, x_{j+1}]$. The linear interpolation formula

$$y = Ay_j + By_{j+1} \quad (\text{A.18})$$

where

$$A = \frac{x_{j+1} - x}{x_{j+1} - x_j} \quad B = 1 - A = \frac{x - x_{j+1}}{x_{j+1} - x_j}, \quad (\text{A.19})$$

satisfies the boundary conditions at the endpoints: $y(x_j) = y_j$ and $y(x_{j+1}) = y_{j+1}$. The problem with this approach is that over the entire range of the function, $[x_1, x_N]$, the result of using Eq. (A.18) would be a series of piecewise linear functions with a zero second derivative within each of the intervals and an undefined or infinite second derivative at the interval boundaries, *i.e.* the grid points $\{x_i\}$. Ideally we would like to construct an interpolation function with a smooth first derivative and a continuous second derivative, both within the intervals and at the the grid points.

By adding a cubic polynomial to the linear interpolation equation, Eq. A.18, within each interval, it is possible to construct an interpolation function that varies linearly in the second derivative. Assume for a moment that we know the values of the second derivative at each grid point: $y''_i = d^2y(x_i)/dx^2, i = 1 \dots N$. We can construct a cubic polynomial that not only has the correct second derivatives at the end points of each interval, y''_j and y''_{j+1} , but also evaluates to zero at those same endpoints. The reason the cubic polynomial must be zero at the endpoints is to not spoil the agreement that is already built into the linear function. A function constructed from these principles is given by

$$y = Ay_j + By_{j+1} + Cy''_j + Dy''_{j+1} \quad (\text{A.20})$$

where

$$C = \frac{1}{6}(A^3 - A)(x_{j+1} - x_j)^2 \quad D = \frac{1}{6}(B^3 - B)(x_{j+1} - x_j)^2. \quad (\text{A.21})$$

To explicitly check that this function does indeed satisfy the conditions at the endpoints, take the derivatives

$$\frac{dy}{dx} = \frac{y_{j+1} - y_j}{x_{j+1} - x_j} - \frac{3A^2 - 1}{6}(x_{j+1} - x_j)y''_j + \frac{3B^2 - 1}{6}(x_{j+1} - x_j)y''_{j+1} \quad (\text{A.22})$$

and

$$\frac{d^2y}{dx^2} = Ay_j'' + By_{j+1}''. \quad (\text{A.23})$$

The second derivative is continuous across the boundary between two intervals, *i.e.* $[x_{j-1}, x_j]$ and $[x_j, x_{j+1}]$, and obeys the conditions at the endpoints since at $x = x_j$, $(A = 1, B = 0)$ and at $x = x_{j+1}$, $(A = 0, B = 1)$.

We had made the assumption that the values of the second derivative are known at the grid points, which of course they are not. By imposing the condition that the first derivative is smooth and continuous across the boundary between two intervals, it is possible to derive a set of equations to generate the y_i'' 's. Evaluate Eq. A.22 at $x = x_j$ for the interval $[x_{j-1}, x_j]$, and set it equal to the same equation evaluated at $x = x_j$ in the interval $[x_j, x_{j+1}]$; rearranging the terms we arrive at

$$\frac{x_j - x_{j+1}}{6}y_{j+1}'' + \frac{x_{j+1} - x_{j-1}}{3}y_j'' + \frac{x_{j+1} - x_j}{6}y_{j+1}'' = \frac{y_{j+1} - y_j}{x_{j+1} - x_j} - \frac{y_j - y_{j+1}}{x_j - x_{j+1}}, \quad (\text{A.24})$$

where $j = 2 \dots N - 1$. To generate a unique solution for the system of $N - 2$ equations, we have to impose boundary conditions at x_1 and x_N . The possibilities being either to set y_1'' and y_N'' to zero, the *natural* cubic spline, or, if you want to make the first derivative at the boundaries to have a specified value, use y_1' and y_N' to calculate the second derivatives at the endpoints using equation. For an excellent discussion of one-dimensional cubic splines see Ref. [189].

A.4 The Runge-Kutta Algorithm

The second order Runge-Kutta method [257] can be used to solve first order differential equations of the form

$$\frac{dy}{dx} = f(x), \quad (\text{A.25})$$

where the solution $y(x)$ is the integral of $f(x)$. The integral equation equivalent to Eq. (A.25) is

$$y(x) = \int_{x_0}^x dx' f(x') + y_0. \quad (\text{A.26})$$

To derive the second order Runge-Kutta method let $y(x)$ and $f(x)$ be discrete functions on a uniform grid x of spacing $h : x_n = x_0 + nh (n = 0, 1, 2, \dots)$. Taylor expand $y(x)$ around x_n to yield

$$y_{n+1} = y_n + hy'_n + \frac{1}{2!}h^2y''_n + \frac{1}{3!}h^3y'''_n + \mathcal{O}[h^4], \quad (\text{A.27})$$

which by using Eq. (A.25) becomes

$$y_{n+1} = y_n + hf_n + \frac{1}{2!}h^2f'_n + \frac{1}{3!}h^3f''_n + \mathcal{O}[h^4]. \quad (\text{A.28})$$

Using the formula for the first derivative

$$f'_n = \frac{f_{n+1} - f_n}{h} + \mathcal{O}[h^2], \quad (\text{A.29})$$

Eq. (A.28) can be written, up to a third order approximation, as

$$y_{n+1} = y_n + h \left[\frac{1}{2}f_n + \frac{1}{2}f_{n+1} \right] + \mathcal{O}[h^3]. \quad (\text{A.30})$$

The algorithm generates the sequence y_n , with the boundary condition $y(x_0) = y_0$, according to the following rules:

$$k_1 = hf_n \quad (\text{A.31})$$

$$k_2 = hf_{n+1} \quad (\text{A.32})$$

$$y_{n+1} = y_n + \frac{1}{2}(k_1 + k_2). \quad (\text{A.33})$$

A.5 The Numerov Algorithm

The Numerov algorithm [257] is used for integrating second-order differential equations of the form

$$\frac{d^2 y}{dx^2} + k^2(x)y = S(x). \quad (\text{A.34})$$

To derive the Numerov algorithm Taylor expand the function $y(x)$ around $x = 0$,

$$y(x) = y_0 + xy' + \frac{1}{2!}x^2 y'' + \frac{1}{3!}x^3 y''' + \mathcal{O}[x^4]. \quad (\text{A.35})$$

Let $y(x)$ be a discrete function on a uniform grid x of spacing $h : x_n = x_0 + nh (n = 0, 1, 2, \dots)$.

Now it is possible to write $y_{n\pm 1}$ in terms of y_n and the derivatives of $y(x)$ evaluated at x_n ,

$$y_{n\pm 1} = y(x = x_n \pm h) = y_n \pm hy'_n + \frac{1}{2!}h^2 y''_n \pm \frac{1}{3!}h^3 y'''_n + \frac{1}{4!}h^4 y''''_n \pm \frac{1}{5!}h^5 y_n^{(5)} + \mathcal{O}[h^6]. \quad (\text{A.36})$$

The equation for the second derivative of $y(x)$ at $x = x_0 + nh$ follows by adding y_{n+1} and y_{n-1} ,

$$\frac{(y_{n+1} - 2y_n + y_{n-1}))}{h^2} = y'' + \frac{1}{12}h^2 y'''' + \mathcal{O}[h^4]; \quad (\text{A.37})$$

the $\mathcal{O}[h^2]$ error has been explicitly written for reasons that will become clear shortly. From the differential equation in Eq. (A.34), it is possible to derive

$$\begin{aligned} y_n'''' &= \frac{d^2}{dx^2} (-k^2 y + S)|_{x=x_n} \\ &= - \left\{ \frac{(k^2 y)_{n+1} - 2(k^2 y)_n + (k^2 y)_{n-1}}{h^2} \right\} + \left\{ \frac{S_{n+1} - 2S_n + S_{n-1}}{h^2} \right\} + \mathcal{O}[h^2]. \end{aligned} \quad (\text{A.38})$$

Now substitute Eq. (A.37) into Eq. (A.39) and rearrange the terms to produce the Numerov algorithm:

$$\begin{aligned} (1 + \frac{h^2}{12}k_{n+2}^2)y_{n+2} - 2(1 - \frac{5h^2}{12}k_{n+1}^2)y_{n+1} + (1 + \frac{h^2}{12}k_n^2)y_n = \\ \frac{h^2}{12}(S_{n+2} + 10S_{n+1} + S_n) + \mathcal{O}[h^6]. \end{aligned} \quad (\text{A.39})$$

The necessary initial conditions are for the first two values of y :

$$y(x_0) = y_0 \quad \text{and} \quad y(x_1) = y_1, \quad (\text{A.40})$$

to start the integration.

A.5.1 Application to the Radial Schrodinger Equation and the Need for Logarithmic Grids

The radial Schrodinger equation

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V(r) \right] u_{nl}(r) = \epsilon_{nl} u_{nl}(r). \quad (\text{A.41})$$

can be represented in a form in Eq. (A.34) with $S(r) = 0$ and

$$k^2(r) = 2[\epsilon_{nl} - \frac{l(l+1)}{r^2} - V(r)]. \quad (\text{A.42})$$

Eq. (A.41) can be solved for bound states ($\epsilon_{nl} < 0$) with the boundary conditions

$$\begin{aligned} \lim_{r \rightarrow 0} u_{nl}(r) &\longrightarrow r(1 - Zr), \quad l = 0 \\ \lim_{r \rightarrow 0} u_{nl}(r) &\longrightarrow r^{l+1}, \quad l > 0 \end{aligned} \quad (\text{A.43})$$

and $u_{nl}(r) \longrightarrow 0$ for $r \rightarrow \infty$. For the case of a bare-core atomic potential $V(r) = -Z/r$, which has a strong cusp at the origin, it is necessary to use a grid with a high density of points near the origin. For this reason we use a logarithmic grid $y = \log(r)$, which for a given number of grid points N has a much higher concentration of points than the linear grid near the origin. Since the Numerov algorithm only applies to uniform grids it is necessary to make the following change of variables

$$y = \log(r) \quad \text{and} \quad Z_{nl}(r) = \frac{u_{nl}(r)}{\sqrt{r}} \quad (\text{A.44})$$

which leads to the modified Schrodinger equation

$$\frac{d^2 Z_{nl}(y)}{dy^2} + k'^2(r) Z_{nl}(y) = 0, \quad (\text{A.45})$$

with

$$k'^2(r) = 2r^2[\epsilon_{nl} - \frac{(l + \frac{1}{2})^2}{2r^2} - V(r)], \quad (\text{A.46})$$

and the boundary conditions

$$\begin{aligned} \lim_{r \rightarrow 0} Z_{nl}(r) &\longrightarrow \sqrt{r}(1 - Zr), \quad l = 0 \\ \lim_{r \rightarrow 0} Z_{nl}(r) &\longrightarrow r^{l+\frac{1}{2}}, \quad l > 0. \end{aligned} \quad (\text{A.47})$$

To solve the Numerov algorithm it is necessary to make an initial guess for the eigenvalue ϵ_{nl}^0 , where the subscript 0 represents the zeroth iteration. It is necessary to iterate until the difference between $\epsilon_{nl}^i - \epsilon_{nl}^{i-1}$ has reached a specified tolerance, usually around $(10^{-9} - 10^{-12})$. For the solution to be acceptable it must have the correct number of nodes and the boundary condition at $r = \infty$ is satisfied. The eigenvalue is adjusted for each iteration by specifying a lower and upper bound, with the eigenvalue chosen to be the average. It is also necessary to determine the classical turning point, where $k'^2(r)$ changes sign, all of the nodes

of the solution must be inside this radius. In the classically forbidden region the solution exponentially decays. The boundary condition near the origin are utilized to determine y_0 and y_1 : the Numerov algorithm integrates outward until the solution exponentially grows in the classically forbidden region or crosses the axis to create an extra node. For the case where the integration produces too many (few) nodes, lower (raise) the eigenvalue; for a divergence in the classically forbidden region raise the eigenvalue. The eigenvalue can either be raised (lowered) by raising (lowering) the lower (upper) bound of the eigenvalue.

A.6 Derivation of the Diffusion Part of the Green's Function

The following derivation is given for a one-electron system, but the results are generalized to the N -electron system. The Green's function for a one-electron system with the Hamiltonian $[H = T = -\nabla^2/2]$ is given by the expression

$$G_D(\mathbf{r} \leftarrow \mathbf{r}'; \tau) = \langle \mathbf{r} | e^{-\tau T} | \mathbf{r}' \rangle.$$

Insert two realizations of the unity operator in the momentum space representation

$$1 = \int d\mathbf{k} |\mathbf{k}\rangle \langle \mathbf{k}|, \quad (\text{A.48})$$

and use the fact that the kinetic energy operator T is diagonal in momentum space ($\mathbf{k} = -i\nabla$) to yield

$$G_D(\mathbf{r} \leftarrow \mathbf{r}'; \tau) = \int d\mathbf{k} \langle \mathbf{r} | \mathbf{k} \rangle e^{-\tau k^2/2} \langle \mathbf{k} | \mathbf{r}' \rangle.$$

Now use the identity

$$\langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k} \cdot \mathbf{r}} \quad (\text{A.49})$$

to derive

$$\begin{aligned} G_D(\mathbf{r} \leftarrow \mathbf{r}'; \tau) &= \frac{1}{(2\pi)^3} \int d\mathbf{k} e^{-\tau k^2/2} e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} \\ &= \frac{1}{(2\pi)^2} \int_0^\infty dk k^2 e^{-\tau k^2/2} \int_0^\pi d\theta \sin(\theta) e^{ik|\mathbf{r} - \mathbf{r}'| \cos(\theta)}. \end{aligned}$$

Using the integral relations

$$\int_0^\pi d\theta \sin(\theta) e^{a \cos(\theta)} = \frac{2 \sinh(a)}{a} \quad (\text{A.50})$$

and

$$\int_0^\infty dr r e^{-ar^2} \sin(rb) = \frac{b e^{-b^2/4a} \sqrt{\pi}}{4a^{3/2}} \quad (\text{A.51})$$

we arrive at the final result

$$G_D(\mathbf{r} \leftarrow \mathbf{r}'; \tau) = \frac{1}{(2\pi\tau)^{3/2}} e^{-(\mathbf{r} - \mathbf{r}')^2/2\tau}.$$

The Green's function for the N -electron system is the product of N one-electron Green's functions:

$$G_D(\mathbf{R} \leftarrow \mathbf{R}'; \tau) = \frac{1}{(2\pi\tau)^{3N/2}} e^{-(\mathbf{R} - \mathbf{R}')^2/2\tau}.$$

The Green's function for the closely related Fokker-Planck equation can be derived following an almost identical procedure as was used for the diffusion equation. The Green's function for the Fokker-Planck equation is given by:

$$G_{FP}(\mathbf{r} \leftarrow \mathbf{r}'; \tau) = \langle \mathbf{r} | e^{\lambda\tau[\nabla^2 - \nabla \cdot \mathbf{F} - \mathbf{F} \cdot \nabla]} | \mathbf{r}' \rangle. \quad (\text{A.52})$$

As a first step insert two realizations of the unity operator in the momentum space representation to yield

$$G_{FP}(\mathbf{r} \leftarrow \mathbf{r}'; \tau) = \int d\mathbf{k} \langle \mathbf{r} | \mathbf{k} \rangle e^{-\lambda\tau[k^2 + 2i\mathbf{k} \cdot \mathbf{F}]} \langle \mathbf{k} | \mathbf{r}' \rangle. \quad (\text{A.53})$$

It is assumed that the function \mathbf{F} is constant over the short time interval τ . Use the identity in Eq. A.49

$$\begin{aligned} G_{FP}(\mathbf{r} \leftarrow \mathbf{r}'; \tau) &= \frac{1}{(2\pi)^3} \int d\mathbf{k} e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} e^{-\lambda\tau[k^2 + 2i\mathbf{k} \cdot \mathbf{F}]} \\ &= \frac{1}{(2\pi)^2} \int dk k^2 e^{-\lambda\tau k^2} \int d\theta \sin\theta e^{ik|\mathbf{r} - \mathbf{r}' - 2\lambda\tau\mathbf{F}| \cos\theta}, \end{aligned} \quad (\text{A.54})$$

and the integral relations, Eqs. A.50 and A.51, to arrive at

$$G_{FP}(\mathbf{r} \leftarrow \mathbf{r}'; \tau) = \frac{1}{(4\pi\lambda\tau)^{3/2}} e^{-(\mathbf{r} - \mathbf{r}' - 2\lambda\tau\mathbf{F})^2 / 4\lambda\tau}. \quad (\text{A.55})$$

The Green's function for the N -electron system is:

$$G_{FP}(\mathbf{R} \leftarrow \mathbf{R}'; \tau) = \frac{1}{(4\pi\lambda\tau)^{3/2}} e^{-(\mathbf{R} - \mathbf{R}' - 2\lambda\tau\mathbf{F})^2 / 4\lambda\tau}. \quad (\text{A.56})$$

A.7 The Cusp Conditions

A.7.1 The Nuclear Cusp Condition

The nuclear cusp condition is the direct result of the divergence of the potential $V(r) = -Z/r$ at the origin, *i.e.* $r \rightarrow 0$. To derive the nuclear cusp condition begin with the Schrodinger equation for a Hydrogen-like system of a single electron and a nucleus of charge Z : in atomic units this is given by

$$\left[-\frac{1}{2}\nabla^2 - \frac{Z}{r} \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (\text{A.57})$$

The divergence of the potential energy as the electron approaches the nucleus is canceled by a divergence in the kinetic energy, resulting in the energy E remaining finite.

As a next step decompose $\psi(\mathbf{r})$ into the product of a radial function and a spherical

harmonic, $\psi(\mathbf{r}) = R(r)Y_{lm}(\theta, \phi)$, and substitute into Eq. (A.57) to yield the radial equation

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \frac{2Z}{r} R - \frac{l(l+1)}{r^2} R + 2ER = 0. \quad (\text{A.58})$$

Let $R(r) = r^m \rho(r)$, where $\rho(r)$ is a function that is finite in the limit $r \rightarrow 0$, with the first and second derivatives $R'(r) = mr^{m-1}\rho + r^m \rho'$ and $R''(r) = m(m-1)r^{m-2}\rho + 2mr^{m-1}\rho' + r^m \rho''$, respectively. The terms containing r^{-2} cancel for $m = l$, reducing Eq. (A.58) to

$$\frac{\rho''}{\rho} + \frac{2(l+1)}{r} \frac{\rho'}{\rho} + \frac{2Z}{r} + 2E = 0. \quad (\text{A.59})$$

Canceling the divergences in the r^{-1} terms leads to the desired electron-nuclear cusp condition

$$\left. \frac{\rho'}{\rho} \right|_{r=0} = -\frac{Z}{l+1}. \quad (\text{A.60})$$

In the $r \rightarrow 0$ limit it is possible to approximate the radial part of $\psi(\mathbf{r})$ by $R^l(r) \approx r^l \exp(-Zr/(l+1))$ [7].

A.7.2 The Electron-Electron Cusp Condition

To derive the electron-electron cusp condition begin with the Hamiltonian for a two-electron system

$$H = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (\text{A.61})$$

which can be written in terms of the relative $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and center-of-mass coordinates

$$\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$$

$$H = -\nabla_r^2 - \frac{1}{4}\nabla_R^2 + \frac{1}{r}. \quad (\text{A.62})$$

Following an almost identical process that was used in deriving the electron-nuclear cusp condition, the Schrodinger equation for the relative coordinate is given by the expression

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \frac{1}{r} R - \frac{l(l+1)}{r^2} R + ER = 0. \quad (\text{A.63})$$

Canceling the r^{-2} terms leads to

$$\frac{\rho''}{\rho} + \frac{2(l+1)}{r} \frac{\rho'}{\rho} - \frac{1}{r} + E = 0 \quad (\text{A.64})$$

which yields the cusp condition

$$\left. \frac{\rho'}{\rho} \right|_{r=0} = \frac{1}{2(l+1)}. \quad (\text{A.65})$$

For the singlet ($S = 0$) case, the electrons are in a s ($l = 0$) like state. This is due to the fact that fermions obey the anti-symmetry condition for particle exchange Eq. (6.1) - if the spin part of the wave function is in an anti-symmetric singlet state, the position space part must be symmetric. For the triplet ($S = 1$) case, the electrons are in a p ($l = 1$) like state. These cases can be summarized by the relation

$$\left. \frac{\rho'}{\rho} \right|_{r=0} = \begin{cases} \frac{1}{2} & \text{for opposite spins} \\ \frac{1}{4} & \text{for parallel spins} \end{cases}. \quad (\text{A.66})$$

A.8 Proof of the Relation Between the Norm-Conservation Condition and the Energy Derivative of the Logarithmic Derivative

To prove the relationship in Eq. (7.8), restated here as

$$-2\pi \left[(r\psi)^2 \frac{\partial}{\partial \epsilon} \frac{d}{dr} \ln \psi \right]_{r_c} = 4\pi \int_0^{r_c} dr r^2 \psi^2, \quad (\text{A.67})$$

we follow procedure similar to that of Shirley *et al.* [205]. It is necessary to begin with the single electron, non-relativistic radial Schrodinger equation for a spherically symmetric potential (in atomic units)

$$-\frac{1}{2} \frac{d^2 \phi}{dr^2} + \left[\frac{l(l+1)}{2r^2} + V - \epsilon \right] \phi = 0, \quad (\text{A.68})$$

where $\phi = r\psi$ is the radial orbital and ϵ is the energy corresponding to ψ (not necessarily an eigenvalue). Define the variable X which is the radial logarithmic derivative of ϕ :

$$X(\epsilon, r) = \frac{d}{dr} [\ln \phi(\epsilon, r)]. \quad (\text{A.69})$$

Differentiate Eq. (A.69) with respect to r

$$X'(\epsilon, r) = \frac{d}{dr} \left(\frac{\phi'}{\phi} \right) = \frac{\phi''}{\phi} - \left(\frac{\phi'}{\phi} \right)^2, \quad (\text{A.70})$$

which leads to the relation

$$[X'(\epsilon, r) + X(\epsilon, r)^2] = \frac{\phi''}{\phi}. \quad (\text{A.71})$$

By use of the radial Schrodinger equation, Eq. (A.68), it is possible to substitute for ϕ'' to yield

$$X'(\epsilon, r) + X(\epsilon, r)^2 = \frac{l(l+1)}{r^2} + 2[V - \epsilon]. \quad (\text{A.72})$$

Now differentiate with respect to ϵ

$$\frac{\partial}{\partial \epsilon} X'(\epsilon, r) + 2X(\epsilon, r) \frac{\partial}{\partial \epsilon} X(\epsilon, r) = -2, \quad (\text{A.73})$$

which can be shown to equal

$$\frac{d}{dr} \left[\phi^2 \frac{\partial}{\partial \epsilon} X(\epsilon, r) \right] = -2\phi^2. \quad (\text{A.74})$$

The final step is to integrate

$$\left[\phi^2 \frac{\partial}{\partial \epsilon} X(\epsilon, r) \right]_{r_c} = -2 \int_0^{r_c} dr \phi^2. \quad (\text{A.75})$$

The above relation is equal to Eq. (A.67) by substituting $\phi = r\psi$ and multiplying both sides by -2π .

A.9 Input File for Gaussian 03

This is an example input file for the Ge_2H_6 molecule for the Gaussian 03 package [11]. This input file includes the Ge and H atomic basis sets and the Gaussian representation of the HF pseudopotential from Ref. [12].

```
%mem=2000000
```

```
%chk=Ge2H6a.chk
```

```
#P HF Gen Pseudo=Read pop=regular units=au
```

Ge2H6 relativistic pseudopotential

O 1

Ge	0.000000000000	0.000000000000	2.27050594990
Ge	0.000000000000	0.000000000000	-2.27050594990
H	0.000000000000	2.690399997713	3.38490611634
H	-2.329954744198	-1.345199998856	3.38490611634
H	2.329954744198	-1.345199998856	3.38490611634
H	0.000000000000	-2.690399997713	-3.38490611634
H	-2.329954744198	1.345199998856	-3.38490611634
H	2.329954744198	1.345199998856	-3.38490611634

Ge 0

SP 1 1.0

1.8893	1.0	1.0
--------	-----	-----

SP 1 1.0

1.2528	1.0	1.0
--------	-----	-----

SP 1 1.0

0.3473	1.0	1.0
--------	-----	-----

SP 1 1.0

0.1199	1.0	1.0
--------	-----	-----

D 1 1.0

0.3225	1.0	1.0
--------	-----	-----

H 0

S 10 1.0

```

5.194042250000e+02 5.993969000000e-04
1.164538900000e+02 1.182210200000e-03
5.929693250000e+01 4.463008800000e-03
3.038276170000e+01 1.965309700000e-03
1.527814390000e+01 2.432256750000e-02
5.228406600000e+00 7.729528480000e-02
2.011882860000e+00 1.919010140000e-01
8.278748700000e-01 3.582211210000e-01
3.528963300000e-01 3.734942420000e-01
1.519936310000e-01 1.114260400000e-01
S 10 1.0
2.057217150000e+02 5.993969000000e-04
4.612417980000e+01 1.182210200000e-03
2.348588260000e+01 4.463008800000e-03
1.203377550000e+01 1.965309700000e-03
6.051252190000e+00 2.432256750000e-02
2.070827920000e+00 7.729528480000e-02
7.968514160000e-01 1.919010140000e-01
3.278984450000e-01 3.582211210000e-01
1.397725210000e-01 3.734942420000e-01
6.020049310000e-02 1.114260400000e-01
S 1 1.0
1.243000000000e-01 1.000000000000e+00
P 1 1.0
1.083800000000e+00 1.000000000000e+00
****

```

Ge 0

DFPP 2 28

Vd

8

1	1.61064071	4.00000000
2	2.09933155	-1.98792973
2	0.52560138	-1.11113943
3	0.96723061	17.12776241
3	1.02293234	-10.68519957
4	1.46268678	-9.96240393
4	2.74284856	-4.17742874
4	2.07410625	9.38249264

Vs - Vd

8

2	3.72786591	3030.66009545
2	4.84415098	-3025.78448986
3	3.47177005	-145.29433187
3	5.17417165	476.96147327
3	5.86299536	-331.66714140
4	1.25319886	-5.57011669
4	4.34019107	-3355.75676252
4	1.48777601	1.86441160

Vp - Vd

8

2	0.54761787	0.89906275
2	6.19779895	0.92815207
3	2.26197595	-19.85908121

3	1.51819178	74.08955303
3	1.24770700	-54.23047181
4	1.67705941	27.51590356
4	2.53042379	-81.64533076
4	2.93857052	60.37426995

A.10 Data for Germanium

This section outlines the procedure to spin-orbit correct the spectral data for Ge.

- A. Start with the spectral data from Ref. [15]. Only the energy levels relevant to computing the excitation (removal and promotion) energies are included, see Table A.3.
- B. Add 63713.24 cm^{-1} to all energies Ge II; add $(63713.24 + 128521.3) \text{ cm}^{-1}$ to all energies Ge III; add $(63713.24 + 128521.3 + 274693) \text{ cm}^{-1}$ to all energies Ge IV. This creates an overall level scheme, see Table A.4.
- C. Reference all levels relative to a level that does not have a spin-orbit splitting because either the total L or S is zero. A suitable choice is the $4s^2$ level, see Table A.5.
- D. Perform spin-orbit averaging. Note that the averages are computed as J weighted $(2J + 1)$ averages, see Table A.6.
- E. The atomic excitation energies are computed as energy differences. The conversion factor is 8065.544 eV/cm [91].

(a) Energy levels of neutral Ge (Ge I).			
Configuration	Term	J	Level(cm^{-1})
$4s^24p^2$	3P	0	0.0000
		1	557.1341
		2	1409.9609
$4s4p^3$	5S	2	41926.726
Ge II ($^2P_{1/2}$)	Limit		63713.24
(b) Energy levels of singly ionized Ge (Ge II).			
Configuration	Term	J	Level(cm^{-1})
$4s^24p$	2P	1/2	0.000
		3/2	1767.357
$4s4p^2$	4P	1/2	51575.885
		3/2	52290.942
		5/2	53366.738
Ge III (1S_0)	Limit		128521.3
(c) Energy levels of doubly ionized Ge (Ge III).			
Configuration	Term	J	Level(cm^{-1})
$4s^2$	1S	0	0.0000
Ge IV ($^2S_{1/2}$)	Limit		274693
(d) Energy levels of triply ionized Ge (Ge IV).			
Configuration	Term	J	Level(cm^{-1})
$4s$	2S	1/2	0.0000
Ge V (1S_0)	Limit		368720

Table A.3: The energy levels of atomic Ge.

A.11 The Local Pseudopotential for Germanium

A.11.1 Introduction

Previous QMC work on Ge have suggested that a local potential is a good approximation [49, 50, 51]. Evaluating a local potential is much less computationally expensive in QMC than evaluating non-local potential. In addition, for a local potential the potential energy depends on the relative positions of the ions and electrons and the functional form of the potential, in contrast the potential energy of a non-local potential explicitly depends on the trial wave function (see Eq. (7.20)). As discussed in Section 7.6.3, using a non-local pseudopotential

Configuration	Term	J	Level(cm ⁻¹)
$4s^24p^2$	3P	0	0.0000
		1	557.1341
		2	1409.9609
$4s4p^3$	5S	2	41926.726
$4s^24p$	2P	1/2	63713.24
		3/2	65480.597
$4s4p^2$	4P	1/2	115289.125
		3/2	116004.182
		5/2	117079.978
$4s^2$	1S	0	192234.54
$4s$	2S	1/2	466927.54

Table A.4: Overall level scheme of energy levels.

Configuration	Term	J	Level(cm ⁻¹)
$4s^24p^2$	3P	0	-192234.54
		1	-191677.4059
		2	-190824.5791
$4s4p^3$	5S	2	-150307.814
$4s^24p$	2P	1/2	-128521.294
		3/2	-126753.943
$4s4p^2$	4P	1/2	-76945.415
		3/2	-76230.358
		5/2	-75154.562
$4s^2$	1S	0	0.000
$4s$	2S	1/2	274693

Table A.5: Energy levels relative to $4s^2$.

in DMC introduces another layer of approximations and violates the variational principle. For these reasons it is always desirable to avoid using non-local pseudopotentials in QMC whenever possible. Therefore, as part of the work for this thesis a local pseudopotential was generated for Ge.

A.11.2 The Generation of the Pseudopotential

All of the HF results in this section are from the numerical HF code discussed in Chapter 2. A modification to this code was used to create the local potential. The advantage to using this code as opposed to the Gaussian 03 package [11] is that there are no basis

Configuration	Term	Level(cm ⁻¹)
4s ² 4p ²	³ P	-191265.517
4s4p ³	⁵ S	-150307.814
4s ² 4p	² P	-127343.06
4s4p ²	⁴ P	-75811.636
4s ²	¹ S	0.000
4s	² S	274693

Table A.6: Spin-orbit averaged energy levels.

set errors and it is possible to explicitly occupy individual orbitals with quantum numbers (n, l, m, s) . All of the atomic configurations used in this study are occupied according to Hund's rules.¹ For example, the (³P) ground state of Ge is created by occupying the $(4, 0, 0, \uparrow), (4, 0, 0, \downarrow), (4, 1, 1, \uparrow), (4, 1, 0, \uparrow)$ atomic orbitals: the (⁵F) excited state is created by occupying the $(4, 0, 0, \uparrow), (4, 1, 1, \uparrow), (4, 1, 0, \uparrow), (4, 2, 2, \uparrow)$ atomic orbitals.

The specific form of the local potential is the Starkloff-Joannopoulos [52] form

$$V_{SJ}(r) = -\frac{Z}{r} \frac{(1 - e^{-ar})}{(1 + e^{a(r_c - r)})}, \quad (\text{A.76})$$

where $Z = 4$ for Ge and a and r_c are adjustable parameters. This potential was found to be a reasonable approximation for Ge in a study by Rajagopal *et al.* [49, 50]. Test using LDA indicated that this potential was significantly inferior to fully norm-conserving non-local potentials but of similar quality to the best pseudo-Hamiltonian [50].

The pseudopotential was constructed to be a norm-conserving pseudopotential. The fitting procedure closely follows the HSC criteria in Section 7.2.1: at the matching radius the partial-norms, eigenvalues and radial logarithmic derivatives of the pseudo orbitals are set to be as closely equal to the corresponding results for the all-electron orbitals. Since the potential is local and is simultaneously fit for the s and p channels, it is difficult to match the quantities with the same level of accuracy as a non-local potential, which have much more variational freedom. The best fit was obtained for the matching radius of 1.7

¹For an excellent reference see Ref. [258].

(nl) Value	Pseudo HF	All-Electron HF
4s	-0.54972	-0.56670
4p	-0.28710	-0.28614

Table A.7: Results from the generation of the local potential: comparison of the eigenvalues from pseudo Hartree-Fock (HF) and all-electron HF.

(nl) Value	Local	Non-Local
4s	-0.54972	-0.56657
4p	-0.28710	-0.28629
Total Energy	-3.63041	-3.66673

Table A.8: Comparison of the eigenvalues from the local pseudopotential to the non-local pseudopotential discussed in Section 7.4

Bohr, the optimal values of the parameters are a and r_c . Table A.7 summarizes the results for the fitting of the pseudopotential. Table A.8 compares the results for the eigenvalues and total energy of the local potential to the non-local potential of Ref. [12, 209]. The non-local pseudopotential (see Section 7.4 for a detailed discussion) is used for almost all of the calculations presented in this thesis. The local potential underbinds the 4s and slightly overbinds the 4p electrons relative to the non-local potential.

A.11.3 Results for Atomic Excitations

We have performed an almost identical set of calculations for the atomic excitations of Ge for the local potential as we did for the non-local potential in Chapter 8. Table A.10 presents the results for the first four ionization energies, the s removal energy, the $s \rightarrow p$ promotion energy and the $s \rightarrow d$ promotion energy. We have also included GW results of Shirley [17] wherever available. The results for the local potential, with the exception of the first two ionization energies, are not as accurate as the results for the non-local potential. This is especially true for the third ionization energy and the s removal energy and $s \rightarrow p$ promotion energies.

Table A.11 contains the total energies for the different atomic configurations of Ge. We have included the results from HF and both VMC and DMC, with and without the CPP. It is

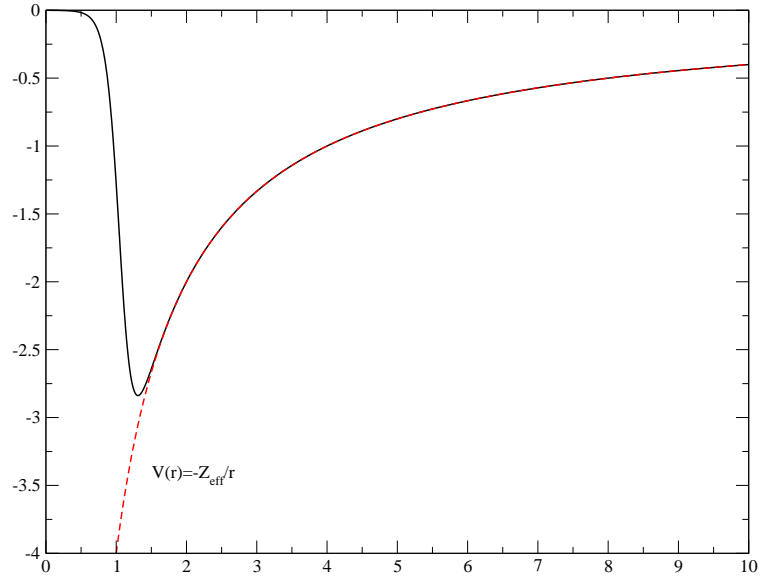


Figure A.1: The local Ge pseudopotential. All units in a.u.

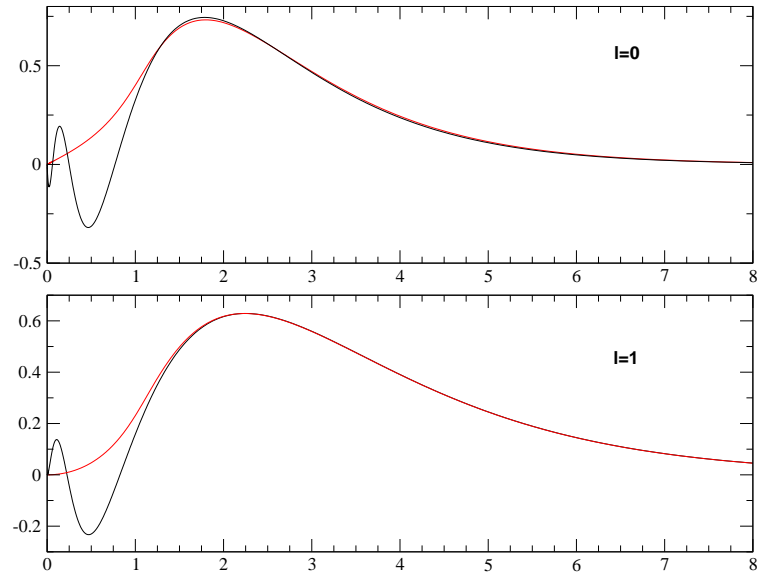


Figure A.2: Plot of the pseudo and all-electron 4s and 4p Ge atomic orbitals. All units in a.u.

reassuring to note that the size of the CPP estimators for the various atomic configurations are almost identical. Therefore, it is not surprising that the size of the CPP corrections for the atomic excitations are also almost identical.

Table A.9: The CPP estimator $\langle \mathcal{H}_{CPP} \rangle$ of various atomic configurations of Ge. All energies in eV.

No. of electrons	Configuration	$\langle \mathcal{H}_{CPP} \rangle$	
		VMC	DMC
4	$4s^2 4p^2(^3P)$	-1.74153(21)	-1.76390(65)
	$4s 4p^3(^5S)$	-1.36199(27)	-1.37570(49)
3	$4s^2 4p(^2P)$	-1.69701(33)	-1.72689(65)
	$4s 4p^2(^4P)$	-1.28790(23)	-1.31023(49)
2	$4s^2(^1S)$	-1.69059(30)	-1.72757(82)
1	$4s(^2S)$	-1.10013(14)	-1.11483(41)

Table A.10: For the Ge atom results for the $s \rightarrow p$ promotion energy, the s removal energy (s^-), the first four ionization energies and the total energy (TE) of the pseudoatom, with and without the CPP. All excitation energies are total energy differences. The CPP correction, ΔCPP , is calculated as the energy difference of the results with and without the CPP. Experimental results from Ref. [15] have been properly spin-orbit corrected [16]: the J weighted averaging has been performed with respect to the $4s^2(^1S)$ configuration [20]. All energies in eV and the QMC error bars are smaller 0.01 eV.

	$s \rightarrow p$ $s^2p^2(^3P) \rightarrow$ $sp^3(^5S)$	s^- $s^2p^2(^3P) \rightarrow$ $sp^2(^4P)$	$s^2p^2(^3P) \rightarrow$ $sp^2d(^5F)$
Expt.	5.08	14.31	-
DMC+CPP	4.71	13.94	11.98
DMC	4.35	13.52	11.55
ΔCPP	0.36	0.42	0.43
VMC+CPP	4.55	13.76	11.84
VMC	4.16	13.30	11.36
ΔCPP	0.39	0.46	0.48
HF	2.94	11.78	9.85

	1st IP $s^2p^2(^3P) \rightarrow$ $s^2p(^2P)$	2nd IP $s^2p(^2P) \rightarrow$ $s^2(^1S)$	3rd IP $s^2(^1S) \rightarrow$ $s(^2S)$	4th IP $s(^2S) \rightarrow$ [Ar] $3d^{10}(^1S)$	TE of $s^2p^2(^3P)$
Expt.	7.93	15.79	34.06	45.72	103.50
DMC+CPP	7.91	15.68	33.57	45.53	102.69
DMC	7.90	15.68	32.98	44.41	100.97
ΔCPP	0.01	0.00	0.59	1.12	1.72
VMC+CPP	7.84	15.59	33.52	45.53	102.48
VMC	7.78	15.59	32.95	44.41	100.73
ΔCPP	0.06	0.00	0.57	1.12	1.75
HF	7.43	15.16	31.78	44.41	98.78

	HF	VMC	VMC+CPP	DMC	DMC+CPP
$(4s)^2(4p)^2$	-98.78840	-100.72243(38)	-102.47892(65)	-100.9702(14)	-102.6898(13)
$(4s)^2(4p)$	-91.35561	-92.94555(27)	-94.64231(49)	-93.0715(10)	-94.7775(11)
$(4s)(4p)^3$	-95.84851	-96.56332(23)	-97.92452(35)	-96.62191(65)	-97.98107(63)
$(4s)(4p)^2$	-87.00542	-87.42627(13)	-88.71907(27)	-87.45009(38)	-88.75338(52)
$(4s)(4p)^2(4d)$	-88.94278	-89.36128(20)	-90.63761(41)	-89.41582(63)	-90.70939(87)
$(4s)^2$	-76.19153	-77.35970(25)	-79.05037(30)	-77.39284(57)	-79.09763(73)
$(4s)$	-44.40898	-44.40898	-45.52550	-44.40898	-45.52550

Table A.11: Hartree-Fock (HF) and QMC total energies for various atomic configurations of Ge. For the $(4s)$ configuration, we did not perform QMC since the numerical Hartree-Fock results are exact for a single electron in a central potential. All energies in eV.

References

- [1] A. Tsolakidis and R. Martin. *Phys. Rev. B* **71**, 125319 (2005).
- [2] A. Williamson, J. Grossman, R. Hood, A. Puzder, and G. Galli. *Phys. Rev. Lett.* **89**, 196803 (2002).
- [3] qmcPACK. <http://www.mcc.uiuc.edu/qmcpack/>.
- [4] C. Fischer. *The Hartree-Fock Method for Atoms: A Numerical Approach* (John Wiley and Sons, New York, 1977).
- [5] E. Clementi and C. Roetti. *Atomic Data and Nuclear Data Tables* **14**, 177 (1974).
- [6] S. Aung, R. Pitzer, and S. Chan. *J. Chem. Phys.* **49**, 2071 (1968).
- [7] B. Hammond, W. Lester, and P. Reynolds. *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, London, 1994).
- [8] C. Umrigar, M. Nightingale, and K. Runge. *J. Chem. Phys.* **99**, 2865 (1993).
- [9] E. Shirley, X. Zhu, and S. Louie. *Phys. Rev B.* **56**, 6648 (1997).
- [10] O. Madelung (editor). *Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology*, volume 17A, chapter Crystal and Solid State Physics (Springer, Berlin, 1984).
- [11] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. “Gaussian 03, Revision C.02.” Gaussian, Inc., Wallingford, CT, 2004.

- [12] J. Trail and R. Needs. *J. Chem. Phys.* **122**, 174109 (2005).
- [13] A. Bergner, M. Dolg, W. Kuechle, H. Stoll, and H. Preuss. *Molec. Phys.* **80**, 1431 (1993).
- [14] “Emsl Gaussian Basis Set Order Form.” <http://www.emsl.pnl.gov/forms/basisform.html>. Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 02/25/04, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact Karen Schuchardt for further information.
- [15] J. Sugar and A. Musgrove. *J. Phys. Chem Ref. Data* **22**, 1213 (1993).
- [16] J. Slater. *Quantum Theory of Atomic Structure*, volume 1 (McGraw-Hill Book Company Inc., New York, 1960).
- [17] E. Shirley. Ph.D. thesis, University of Illinois Urbana-Champaign (1991).
- [18] B. Pullen, T. Carlson, W. Moddeman, G. Schweitzer, W. Bull, and F. Grimm. *J. Chem. Phys.* **53**, 768 (1970).
- [19] U. Itoh, Y. Toyoshima, H. Onuki, N. Washida, and T. Ibuki. *J. Chem. Phys.* **85**, 4876 (1986).
- [20] E. Shirley. “Private communication.”
- [21] E. Degoli, G. Cantele, E. Luppi, R. Magri, D. Ninno, O. Bisi, and S. Ossicini. *Phys. Rev. B* **69**, 155411 (2004).
- [22] S. Takeoka, M. Fujii, S. Hayashi, and K. Yamamoto. *Phys. Rev. B* **58**, 7921 (1998).
- [23] S. Ögüt, J. Chelikowsky, and S. Louie. *Phys. Rev. Lett.* **79**, 1770 (1997).
- [24] I. Vasiliev, S. Ögüt, and J. Chelikowsky. *Phys. Rev. Lett.* **86**, 1813 (2001).
- [25] M. Rohlfing and S. Louie. *Phys. Rev. Lett.* **80**, 3320 (1998).
- [26] A. Porter, M. Towler, and R. Needs. *Phys. Rev. B* **64**, 035320 (2001).
- [27] F. Fehér. *Molekülspektroskopische Untersuchungen auf dem Gebiet der Silane und der Heterocyclischen Sulfane, Forschungsbericht des Landes Nordrhein-Westfalen* (West-Deutscher Verlag, Köln, 1977).
- [28] G. Neshet, L. Kronik, and J. Chelikowsky. *Phys. Rev. B* **71**, 035344 (2005).

- [29] R. Martin. *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).
- [30] P. Reynolds, D. Ceperley, B. Alder, and W. Lester. *J. Chem. Phys.* **77**, 5593 (1982).
- [31] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal. *Rev. Mod. Phys.* **73**, 33 (2001).
- [32] G. Bachelet and N. Christensen. *Phys. Rev. B* **31**, 879 (1985).
- [33] D. Ceperley. *J. of Stat. Phys.* **43**, 815 (1986).
- [34] G. Bachelet, D. Ceperley, and M. Chiochetti. *Phys. Rev. Lett.* **62**, 2088 (1989).
- [35] W. Foulkes and M. Schluter. *Phys. Rev. B* **42**, 11505 (1990).
- [36] A. Ma, N. Drummond, M. Towler, and R. Needs. *Phys. Rev. E* **71**, 066704 (2005).
- [37] W. Aulbur, L. Jönsson, and J. Wilkens. *Solid State Physics* **54**, 1 (2000).
- [38] M. Hybersten and S. Louie. *Phys. Rev. B* **34**, 5390 (1986).
- [39] M. Hybersten and S. Louie. *Phys. Rev. Lett.* **56**, 1418 (1985).
- [40] W. Müller, J. Flesch, and W. Meyer. *J. Chem. Phys.* **80**, 3297 (1984).
- [41] W. Müller and W. Meyer. *J. Chem. Phys.* **80**, 3311 (1984).
- [42] G. Jeung, J. Malrieu, and J. Daudey. *J. Chem. Phys.* **77**, 3571 (1982).
- [43] E. Shirley and R. Martin. *Phys. Rev. B* **47**, 15413 (1993).
- [44] E. Shirley, L. Mitas, and R. Martin. *Phys. Rev. B* **44**, 3395 (1991).
- [45] R. Maezono, M. Towler, Y. Lee, and R. Needs. *Phys. Rev. B* **68**, 165103 (2003).
- [46] D. Alfè, M. Gillan, M. Towler, and R. Needs. *Phys. Rev. B* **70**, 214102 (2004).
- [47] Y. Lee and R. Needs. *Phys. Rev. B* **67**, 035121 (2003).
- [48] A. Porter, O. Al-Mushadani, M. Towler, and R. Needs. *J. Chem. Phys.* **114**, 7795 (2001).
- [49] G. Rajagopal, R. Needs, S. Kenny, W. Foulkes, and A. James. *Phys. Rev. Lett.* **73**, 1959 (1994).
- [50] G. Rajagopal, R. Needs, A. James, S. Kenny, and W. Foulkes. *Phys. Rev. B* **51**, 10591 (1995).
- [51] A. Williamson, S. Kenny, G. Rajagopal, A. James, R. Needs, L. Fraser, W. Foulkes, and P. Maccullum. *Phys. Rev. B* **53**, 9640 (1996).
- [52] T. Starkloff and J. Joannopoulos. *Phys. Rev. B* **16**, 5212 (1977).

- [53] J. Wilcoxon, P. Provencio, and G. Samara. *Phys. Rev. B* **64**, 035417 (2001).
- [54] G. Taraschi, S. Saini, W. Fan, L. Kimerling, and E. Fitzgerald. *J. Appl. Phys.* **93**, 9988 (2003).
- [55] A. Konchenko, Y. Nakayama, I. Matsuda, S. Hasegawa, Y. Nakamura, and M. Ichikawa. *Phys. Rev. B* **73**, 113311 (2006).
- [56] L. Benedict, A. Puzder, A. Williamson, J. Grossman, G. Galli, J. Klepeis, J.-Y. Raty, and O. Pankratov. *Phys. Rev. B* **68**, 085310 (2003).
- [57] J. Grossman, M. Rohlfing, L. Mitas, S. Louie, and M. Cohen. *Phys. Rev. Lett.* **86**, 472 (2001).
- [58] Y. Niquet, G. Allan, C. Delerue, and M. Lannoo. *Appl. Phys. Lett.* **77**, 1182 (2000).
- [59] F. Reboredo and A. Zunger. *Phys. Rev. B* **62**, 2275 (2000).
- [60] F. Reboredo and A. Zunger. *Phys. Rev. B* **63**, 235314 (2001).
- [61] D. Melnikov and J. Chelikowsky. *Solid State Commun.* **127**, 361 (2003).
- [62] D. Melnikov and J. Chelikowsky. *Phys. Rev. B* **69**, 113305 (2004).
- [63] H.-C. Weissker, J. Furthmüller, and F. Bechstedt. *Phys. Rev. B* **65**, 155328 (2002).
- [64] H.-C. Weissker, J. Furthmüller, and F. Bechstedt. *Phys. Rev. B* **67**, 245304 (2003).
- [65] H.-C. Weissker, J. Furthmüller, and F. Bechstedt. *Phys. Rev. B* **69**, 115310 (2004).
- [66] A. Puzder, A. Williamson, J. Grossman, and G. Galli. *Phys. Rev. Lett.* **88**, 097401 (2002).
- [67] L. Mitas, J. Therrien, R. Twisten, G. Belomoin, and M. Nayfeh. *Appl. Phys. Lett.* **78**, 1918 (2001).
- [68] M. Wolkin, J. Jorne, P. Fauchet, G. Allan, and C. Delerue. *Phys. Rev. Lett.* **82**, 197 (1999).
- [69] I. Vasiliev, J. Chelikowsky, and R. Martin. *Phys. Rev. B* **65**, 121302 (2002).
- [70] O. Lehtonen and D. Sundholm. *Phys. Rev. B* **72**, 085424 (2005).
- [71] M. Payne, M. Teter, D. Allan, T. Arias, and J. Joannopoulos. *Rev. Mod. Phys.* **64**, 1045 (1992).
- [72] A. Yoffe. *Adv. Phys.* **50**, 1 (2001).
- [73] Y. Kanzawa, T. Kageyama, S. Takeoka, M. Fujii, S. Hayashi, and K. Yamamoto. *Solid State Commun.* **102**, 533 (1997).

- [74] G. Belomoin, J. Therrien, A. Smith, S. Rao, R. Twesten, S. Chaieb, M. H. Nayfeh, L. Wagner, and L. Mitas. *Appl. Phys. Lett.* **80**, 841 (2002).
- [75] Y. Maeda, N. Tsukamoto, Y. Yazawa, Y. Kanemitsu, and Y. Masumoto. *Appl. Phys. Lett.* **59**, 3168 (1991).
- [76] S. Hayashi, M. Fujii, and K. Yamamoto. *Jpn. J. Appl. Phys., Part 2* **28** (1989).
- [77] J. Zhu, C. White, J. Budai, and S. Withrow. *J. Appl. Phys.* **78**, 4386 (1995).
- [78] M. Nayfeh, N. Barry, J. Therrien, O. Akcakir, E. Gratton, and G. Belomoin. *Appl. Phys. Lett.* **78**, 1131 (2001).
- [79] Y. Kanemitsu, H. Uto, Y. Masumoto, and Y. Maeda. *Appl. Phys. Lett.* **61**, 2187 (1992).
- [80] Y. Maeda. *Phys. Rev. B* **51**, 1658 (1995).
- [81] S. Okamoto and Y. Kanemitsu. *Phys. Rev. B* **54**, 16421 (1996).
- [82] D. Paine, C. Caragianis, T. Kim, Y. Shigesato, and T. Ishahara. *Appl. Phys. Lett.* **62**, 2842 (1993).
- [83] A. Dutta. *Appl. Phys. Lett.* **68**, 1189 (1996).
- [84] M. Nogami and Y. Abe. *Appl. Phys. Lett.* **65**, 2545 (1994).
- [85] M. Ludwig, R. Hummel, and S.-S. Chang. *J. Vac. Sci. Technol. B* **12**, 3023 (1994).
- [86] A. Saito and T. Suemoto. *Phys. Rev. B* **56**, 1688 (1997).
- [87] V. Craciun, C. Leborgne, E. Nicholls, and I. Boyd. *Appl. Phys. Lett.* **69**, 1506 (1996).
- [88] M. Zacharias and P. Fauchet. *Appl. Phys. Lett.* **71**, 380 (1997).
- [89] W. Choi, V. Ng, S. Ng, H. Thio, Z. Shen, and W. Li. *J. Appl. Phys.* **86**, 1398 (1999).
- [90] A. Franceschetti and S. Pantelides. *Phys. Rev. B* **68**, 033313 (2003).
- [91] P. Mohr and B. Taylor. “Fundamental Physical Constants from NIST.” <http://physics.nist.gov/constants>.
- [92] R. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- [93] R. Jones and O. Gunnarsson. *Rev. Mod. Phys.* **61**, 689 (1989).
- [94] R. Dreizler and E. Gross. *Density-Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer, Berlin, 1990).
- [95] P. Hohenberg and W. Kohn. *Phys. Rev.* **136**, B864 (1964).

- [96] W. Kohn and L. Sham. *Phys. Rev.* **140**, 1133 (1965).
- [97] D. Ceperley and B. Alder. *Phys. Rev. Lett.* **45**, 566 (1980).
- [98] B. Tanatar and D. Ceperley. *Phys. Rev. B.* **39**, 5005 (1989).
- [99] J. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.* **77**, 3865 (1996).
- [100] J. Perdew and Y. Wang. *Phys. Rev. B* **33**, 8800 (1986).
- [101] J. Perdew. *Phys. Rev. B* **33**, 8822 (1986).
- [102] M. Staedele, J. Majewski, P. Vogl, and A. Gorling. *Phys. Rev. Lett.* **79**, 2089 (1997).
- [103] V. Fock. *Z. Phys.* **61**, 126 (1930).
- [104] J. Slater. *Phys. Rev.* **35**, 210 (1930).
- [105] J. Slater. *Quantum Theory of Atomic Structure*, volume 2 (McGraw-Hill Book Company Inc., New York, 1960).
- [106] J. Thijssen. *Computational Physics* (Cambridge University Press, Cambridge, 1999).
- [107] F. Herman and S. Skillman. *Atomic Structure Calculations* (Prentice-Hall, Engelwood Cliffs, N.J., 1963).
- [108] B. Bransden and C. Joachain. *Physics of Atoms and Molecules* (Addison Wesley Longman Limited, Edinburgh Gate, 1983).
- [109] J. Vincent and J. Kim. “AtomicHF.” <http://www.mcc.uiuc.edu/qmc/AtomicHF/index.html>.
- [110] S. Gasiorowicz. *Quantum Physics* (John Wiley and Sons, Inc., New York, 1996).
- [111] J. Jackson. *Classical Electrodynamics, Second Edition* (John Wiley and Sons, New York, 1975).
- [112] V. Heine. *Group Theory in Quantum Mechanics* (Pergamon Press, London, 1960).
- [113] A. Messiah. *Quantum Mechanics, Vol. II* (John Wiley and Sons, New York, 1964).
- [114] J. Bjorken and S. Drell. *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
- [115] J. Kübler. *Theory of Itinerant Electron Magnetism* (Oxford University Press, New York, 2000).
- [116] D. Koelling and B. Harmon. *J. Phys. C* **10**, 3107 (1977).
- [117] E. Runge and E. Gross. *Phys. Rev. Lett.* **52**, 997 (1984).

- [118] G. Onida, L. Reining, and A. Rubio. *Rev. Mod. Phys.* **74**, 601 (2002).
- [119] A. Tsolakidis. Ph.D. thesis, University of Illinois Urbana-Champaign (2003).
- [120] E. Gross, C. Ullrich, and U. Gossmann. *Density Funcional Theory* (Plenum Press, New York, 1995).
- [121] E. Gross, J. Dobson, and M. Petersilka. *Topics in Current Chemistry*, volume 181 (Springer-Verlag, Berlin Heidelberg, 1996).
- [122] A. Zangwill and P. Soven. *Phys. Rev. Lett.* **45**, 204 (1980).
- [123] A. Tsolakidis, D. Sánchez-Portal, and R. Martin. *Phys. Rev. B* **66**, 235416 (2002).
- [124] S. Rao, J. Sutin, R. Clegg, E. Gratton, M. Nayfeh, S. Habbal, A. Tsolakidis, and R. Martin. *Phys. Rev. B* **69**, 205319 (2004).
- [125] K. Yabana and G. Bertsch. *Phys. Rev. B* **54**, 4484 (1996).
- [126] N. Ashcroft and N. Mermin. *Solid State Physics* (Saunders College, New York, 1976).
- [127] C. Yannouleas and U. Landman. *Eur. Phys. J. D* **16**, 373 (2001).
- [128] C. Yannouleas and U. Landman. *Int. J. Quantum Chem.* **90**, 699 (2002).
- [129] C. Yannouleas and U. Landman. *J. Phys.: Condens. Matter* **14**, 591 (2002).
- [130] H. Heitler and F. London. *Z. Phys.* **44**, 455 (1927).
- [131] P. Ring and P. Schuck. *The Nuclear Many-Body Problem* (Springer-Verlag, New York, 1980).
- [132] M. Hybertsen and S. Louie. *Phys. Rev. B* **37**, 2733 (1988).
- [133] A. Szabo and N. Ostlund. *Modern Quantum Chemistry* (McGraw-Hill Book Company Inc., New York, 1989).
- [134] E. Shirley and R. Martin. *Phys. Rev. B* **47**, 15404 (1993).
- [135] L. Hedin and S. Lundqvist. *Solid State Physics* **23**, 1 (1969).
- [136] A. Fetter and J. Walecka. *Quantum Theory of Many-Particle Systems* (McGraw-Hill Book Company Inc., New York, 1971).
- [137] M. Kalos and P. Whitlock. *Monte Carlo Methods: Volume 1* (John Wiley and Sons, New York, 1986).
- [138] “Boost C++ Libraries.” <http://www.boost.org/>.
- [139] “The Scalable Parallel Random Number Generators Library (SPRNG).” <http://sprng.cs.fsu.edu/>. Version 2.0.

- [140] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 1 (Wiley, New York, 1968), 3rd edition.
- [141] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. *J. Chem. Phys.* **21**, 1087 (1953).
- [142] M. Allen and D. Tildesley. *Computer Simulation of Liquids* (Oxford Science Publications, Oxford, 1987).
- [143] E. Draeger, J. Shumway, and D. Ceperley. “DataSpork.” <http://www.mcc.uiuc.edu/qmc/dataspork.html>. Version 1.1.
- [144] M. Donsker and M. Kac. *J. Res. Natl. Bur. Std.* **44**, 551 (1950).
- [145] W. McMillan. *Phys. Rev. A* **138**, 442 (1965).
- [146] M. Kalos. *Phys. Rev.* **128**, 1791 (1962).
- [147] M. Kalos, D. Levesque, and L. Verlet. *Phys. Rev. A* **9**, 2178 (1974).
- [148] J. Anderson. *J. Chem. Phys.* **63**, 1499 (1975).
- [149] J. Anderson. *J. Chem. Phys.* **65**, 4121 (1976).
- [150] D. Ceperley and G. C. M. Kalos. *Phys. Rev. B* **16**, 3081 (1977).
- [151] D. Ceperley and M. Kalos. *Monte Carlo Methods in Statistical Physics* (Springer, Berlin, 1979), 2nd edition.
- [152] J. Moskowitz, K. Schmidt, M. Lee, and M. Kalos. *J. Chem. Phys.* **77**, 349 (1982).
- [153] S. Fahy, X. Wang, and S. Louie. *Phys. Rev. B* **42**, 3503 (1990).
- [154] J. Shumway and D. Ceperley. “Quantum Monte Carlo Methods in the Study of Nanostructures.” In M. Rieth and W. Schommers (editors), *Handbook of Theoretical and Computational Nanotechnology* (American Scientific Publishers, 2006).
- [155] C. Umrigar, K. Wilson, and J. Wilkens. *Phys. Rev. Lett.* **60**, 1719 (1988).
- [156] Y. Kwon, D. Ceperley, and R. Martin. *Phys. Rev. B* **50**, 1684 (1994).
- [157] Y. Kwon, D. Ceperley, and R. Martin. *Phys. Rev. B* **53**, 7376 (1996).
- [158] M. Kalos. *J. Comput. Phys.* **2**, 257 (1967).
- [159] R. Subramaniam, M. Lee, K. Schmidt, and J. Moskowitz. *J. Chem. Phys.* **97**, 2600 (1988).
- [160] S. Baroni and S. Moroni. *Phys. Rev. Lett.* **82**, 4745 (1999).

- [161] S. Baroni and S. Moroni. “Reptation Quantum Monte Carlo.” In J. Grotendorst, D. Marx, and A. Muramatsu (editors), *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms* (John von Neumann Institute for Computing, Jülich, 2002).
- [162] H. Risken. *The Fokker-Planck Equation* (Springer-Verlag, Berlin, 1984).
- [163] R. Barnett, P. Reynolds, and W. Lester. *J. Comp. Phys.* **96**, 258 (1991).
- [164] M. Kalos. “Private communication.”
- [165] G. Ortiz, D. Ceperley, and R. Martin. *Phys. Rev. Lett.* **71**, 2777 (1993).
- [166] L. Mitas. “Private communication.”
- [167] W. Foulkes, R. Hood, and R. Needs. *Phys. Rev. B* **60**, 4558 (1999).
- [168] M. Towler, R. Hood, and R. Needs. *Phys. Rev. B* **62**, 2330 (2000).
- [169] E. Draeger, J. Grossman, A. Williamson, and G. Galli. *Phys. Rev. Lett.* **90**, 167402 (2003).
- [170] L. Mitas and R. Martin. *Phys. Rev. Lett.* **72**, 2438 (1994).
- [171] L. Mitas. *Comput. Phys. Commun.* **96**, 107 (1996).
- [172] A. Williamson, R. Hood, R. Needs, and G. Rajagopal. *Phys. Rev. B* **57**, 12140 (1998).
- [173] B. Stroustrup. *The C++ Programming Language* (Addison-Wesley, New York, 1997).
- [174] “NCSA HDF5 (Hierarchical Data Format 5) Software Library and Utilities.” <http://www.hdfgroup.com/>.
- [175] W. Gropp, E. Lusk, and A. Skjellum. *Using MPI* (The MIT Press, Cambridge, 1999).
- [176] “GSL - GNU Scientific Library.” <http://www.gnu.org/software/gsl/>.
- [177] “Blitz++.” <http://www.oonumerics.org/blitz/>.
- [178] “SGI - Service and Support: Standard Template Library Programmer’s Guide.” <http://www.sgi.com/tech/stl/>.
- [179] R. Jastrow. *Phys. Rev.* **98**, 1479 (1955).
- [180] K. Riley, M. Hobson, and S. Bence. *Mathematical Methods for Physics and Engineering* (Cambridge University Press, Cambridge, 1998).
- [181] J. Soler, E. Artacho, J. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal. *J. Phys.: Condens. Matter* **14**, 2745 (2002).

- [182] J. Junquera, Ó. Paz, D. Sánchez-Portal, and E. Artacho. *Phys. Rev. B* **64**, 235111 (2001).
- [183] D. Das. Ph.D. thesis, University of Illinois Urbana-Champaign (2005).
- [184] A. Ma, M. Towler, N. Drummond, and R. Needs. *J. Chem. Phys.* **122**, 224322 (2005).
- [185] R. Barnett, P. Reynolds, and W. Lester. *J. Chem. Phys.* **82**, 2700 (1985).
- [186] D. Ceperley. *Phys. Rev. B* **18**, 3126 (1978).
- [187] D. Bressanini, G. Morosi, and M. Mella. *J. Chem. Phys.* **116**, 5345 (2002).
- [188] D. Ceperley and B. Alder. *Phys. Rev. B* **36**, 2092 (1987).
- [189] *Numerical Recipes in C: The Art of Scientific Computing, Second Edition* (Cambridge University Press, Cambridge, 1992).
- [190] M. Casula, C. Attaccalite, and S. Sorella. *J. Chem. Phys.* **121**, 7110 (2004).
- [191] P. Kent. Ph.D. thesis, University of Cambridge (1999).
- [192] P. Kent, R. Needs, and G. Rajagopal. *Phys. Rev. B* **59**, 12344 (1999).
- [193] W. Ku and A. G. Eguiluz. *Phys. Rev. Lett.* **89**, 126401 (2002).
- [194] S. Louie. “Private communication.”
- [195] A. Nicklass and H. Stoll. *Molec. Phys.* **86**, 317 (1995).
- [196] G. Igel-Mann, H. Stoll, and H. Preuss. *Molec. Phys.* **65**, 1321 (1988).
- [197] W. Kuchle, M. Dolg, H. Stoll, and H. Preuss. *Molec. Phys.* **74**, 1245 (1991).
- [198] E. Shirley, R. Martin, G. Bachelet, and D. Ceperley. *Phys. Rev. B* **42**, 5057 (1990).
- [199] D. Hamann. *Phys. Rev. B* **40**, 2980 (1989).
- [200] D. Hamann, M. Schlüter, and C. Chiang. *Phys. Rev. Lett* **43**, 1494 (1979).
- [201] G. Bachelet, D. Hamann, and M. Schlüter. *Phys. Rev. B* **26**, 4199 (1982).
- [202] C. Reis, J. Pacheco, and J. Martins. *Phys. Rev. B* **68**, 155111 (2003).
- [203] W. Pickett. *Computer Physics Reports* **9**, 115 (1988).
- [204] P. Giannozzi. “Notes on Pseudopotential Generation.” <http://web1.sns.it/~giannozz/Atom/doc.pdf>.
- [205] E. Shirley, D. Allan, R. Martin, and J. Joannopoulos. *Phys. Rev. B* **40**, 3652 (1989).
- [206] N. Troullier and J. Martins. *Phys. Rev. B* **43**, 1993 (1991).

- [207] Y. Lee, P. Kent, M. Towler, R. Needs, and G. Rajagopal. *Phys. Rev. B* **62**, 13347 (2000).
- [208] J. Trail and R. Needs. *J. Chem. Phys.* **122**, 014112 (2005).
- [209] “CASINO Pseudopotential Library.” http://www.tcm.phy.cam.ac.uk/~mdt26/pseudo_lib/ge/pseudo.html.
- [210] L. Mitas, E. Shirley, and D. Ceperley. *J. Chem. Phys.* **95**, 3467 (1991).
- [211] M. Hurley and P. Christiansen. *J. Chem. Phys.* **86**, 1069 (1987).
- [212] H. Goldstein. *Classical Mechanics, Second Edition* (Addison-Wesley Publishing House, New York, 1950).
- [213] M. Born and W. Heisenberg. *Z. Phys.* **23**, 388 (1924).
- [214] B. Farid and R. Needs. *Phys. Rev. B* **45**, 1067 (1992).
- [215] S. Chiesa. “Private communication.”
- [216] K. Dyall, P. Taylor, K. Faegri, and H. Partridge. *J. Chem. Phys.* **95**, 2583 (1991). And references therein.
- [217] B. Beagley and J. Monaghan. *Trans. Faraday Soc.* **66**, 2745 (1970).
- [218] U. Steinbrenner, A. Bergner, M. Dolg, and H. Stoll. *Molec. Phys.* **82**, 3 (1994).
- [219] A. Ricca and C. Bauschlicher. *J. Phys. Chem. A* **103**, 11121 (1999).
- [220] A. Kokalj. *Comp. Mater. Sci.* **28**, 155 (2003).
- [221] C. Wu, F. Chen, and D. Judge. *J. Chem. Phys.* **99**, 1530 (1993).
- [222] M. Mitchell, X. Wang, C. Chin, M. Suto, and L. Lee. *J. Phys. B: At. Mol. Phys.* **20**, 5451 (1987).
- [223] L. Chantranupong, G. Hirsch, R. Buenker, and M. Dillon. *Chem. Phys.* **170**, 167 (1993).
- [224] K. Das and K. Balasubramanian. *J. Chem. Phys.* **93**, 5883 (1990).
- [225] T. Kudo and S. Nagase. *Chem. Phys. Lett.* **148**, 73 (1988).
- [226] C. Schubert, U. Kaiser, A. Hedler, W. Wescha, T. Gorelik, U. Glatzel, J. Kräußlich, B. Wunderlich, G. Heß, and K. Goetz. *J. Appl. Phys.* **91**, 1520 (2002).
- [227] T. Buuren, L. Dinh, L. Chase, W. Siekhaus, and L. Terminello. *Phys. Rev. Lett.* **80**, 3803 (1998).
- [228] J. Wilcoxon, G. Samara, and P. Provencio. *Phys. Rev. B* **60**, 2704 (1999).

- [229] D. Lockwood, A. Wang, and B. Bryskiewicz. *Solid State Commun.* **89**, 587 (1994).
- [230] S. Furukawa and T. Miyasato. *Phys. Rev. B* **38**, 5726 (1988).
- [231] H. Takagi, H. Ogawa, Y. Yamazaki, A. Ishizaki, and T. Nakagiri. *Appl. Phys. Lett.* **56**, 24 (1990).
- [232] S. Schuppler, S. Friedman, M. Marcus, D. Adler, Y.-H. Xie, F. Ross, T. Harris, W. Brown, Y. Chabal, L. Brus, and P. Citrin. *Phys. Rev. Lett.* **72**, 2648 (1994).
- [233] S. Schuppler, S. Friedman, M. Marcus, D. Adler, Y.-H. Xie, F. Ross, Y. Chabal, T. Harris, L. Brus, W. Brown, E. Chaban, P. Szajowski, S. Christman, and P. Citrin. *Phys. Rev. B* **52**, 4910 (1995).
- [234] A. Shklyaevev and M. Ichikawa. *Surf. Sci.* **514**, 19 (2002).
- [235] C. Garoufalidis, A. Zdetsis, and S. Grimme. *Phys. Rev. Lett.* **87**, 276402 (2001).
- [236] L.-W. Wang and A. Zunger. *J. Phys. Chem.* **98**, 2158 (1994).
- [237] C. Delerue, G. Allen, and M. Lannoo. *Phys. Rev. B* **48**, 11024 (1993).
- [238] T. Takagahara and K. Takeda. *Phys. Rev. B* **53** (1996).
- [239] N. Hill, S. Pokrant, and A. Hill. *J. Phys. Chem. B* **103**, 3156 (1999).
- [240] M. Palummo, G. Onida, and R. D. Sole. *Phys. Status Solidi A* **175**, 23 (1999).
- [241] A. Becke. *J. Chem Phys.* **98**, 5648 (1993).
- [242] Z. Levine and S. Louie. *Phys. Rev. B* **25**, 6310 (1982).
- [243] P. Yu and M. Cardona. *Fundamentals of Semiconductors* (Springer, Berlin, 2003), 3 edition.
- [244] J. Proot, C. Delerue, and G. Allen. *Appl. Phys. Lett.* **61**, 1948 (1992).
- [245] L. N. Dinh, W. McLean, M. Schilbach, and M. Balooch. *Phys. Rev. B* **59**, 15513 (1999).
- [246] D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. Soler. *Int. J. Quantum Chem.* **65**, 453 (1997).
- [247] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. García, and J. Soler. *Phys. Status Solidi B* **215**, 809 (1999).
- [248] P. Ordejón. *Phys. Status Solidi B* **217**, 335 (2000).
- [249] I. Tokatly and O. Pankratov. *Phys. Rev. Lett.* **86**, 2078 (2001).
- [250] M. Marzari and D. Vanderbilt. *Phys. Rev. B* **56**, 12847 (1997).

- [251] A. Mostoft, J. R. Yates, N. Marzari, I. Souza, and D. Vanderbilt.
<http://www.wannier.org/>.
- [252] A. Williamson, R. Hood, and J. Grossman. *Phys. Rev. Lett.* **87**, 246406 (2001).
- [253] F. Reboredo and A. Williamson. *Phys. Rev. B* **71**, 121105 (2005).
- [254] C.Lin, F.-H. Zong, and D. Ceperley. *Phys. Rev. E* **64**, 016702 (2001).
- [255] V. Natoli and D. Ceperley. *J. Comp. Phys.* **117**, 171 (1995).
- [256] K. Esler. Ph.D. thesis, University of Illinois Urbana-Champaign (2006).
- [257] S. Koonin and D. Meredith. *Computational Physics* (Addison-Wesley Publishing Co., New York, 1990).
- [258] D. Griffiths. *Introduction to Quantum Physics* (Prentice Hall, Englewood Cliffs, NJ, 1995).

Author's Biography

Jordan Eric Vincent was born on December 25, 1976 in Cleveland, OH. He received a B.S. in Physics from The Ohio State University in 1999, and an M.S. in Physics from the University of Illinois at Urbana-Champaign in 2003.